

THE IMPACT OF THE DATA MANAGEMENT APPROACH  
ON INFORMATION SYSTEMS AUDITING

by

DON FRIEDRICH FÜRSTENBURG

submitted in fulfilment of the requirements for  
the degree of

MASTER OF ACCOUNTING SCIENCE

in the

DEPARTMENT OF APPLIED ACCOUNTANCY

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF E SADLER

JOINT SUPERVISOR: PROF C HATTINGH

NOVEMBER 1992

## ACKNOWLEDGEMENTS

I would like to take the opportunity to express my appreciation to those who contributed to the success of this study.

Firstly, I would like to acknowledge the important role that Mr. P.J. de Bruyn and a previous colleague (who wants to remain anonymous) have played. Without them the study would definitely not have progressed beyond the inception stage.

Secondly, to Mr. S. Loubser, who did pioneering work in the field of data resource management in South Africa and who introduced me to the subject of data management.

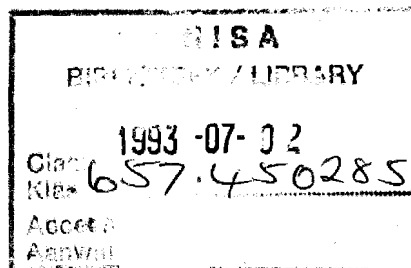
To my supervisor, Professor E. Sadler and joint supervisor, Professor C. Hattingh, I am indebted for their ideas and support and providing feedback when I needed it. They played an invaluable role in facilitating the completion of this research study.

Finally, I thank my wife and children for their understanding, patience and support during the completion of this thesis.

Above all, to God be the honour and glory.

PRETORIA

30 November 1992



## SUMMARY

In establishing the impact of formal data management practices on systems and systems development auditing in the context of a corporate data base environment, the most significant aspects of a data base environment as well as the concept of data management were researched.

It was established that organisations need to introduce a data management function to ensure the availability and integrity of data for the organisation.

It was further established that an effective data management function can fulfil a key role in ensuring the integrity of the overall data base and as such it becomes an important general control on which the auditor can rely.

The audit of information systems in a data base environment requires a more "holistic" audit approach and as a result the auditor has to expand the scope of the systems audit to include an evaluation of the overall data base environment.

## CONTENTS

	Page
<b>LIST OF FIGURES</b>	xiv
<b>LIST OF TABLES</b>	xv
 <b>CHAPTER 1 : OVERVIEW</b>	
 1.1 Introduction	1
 1.2 The concepts of a corporate data base environment	2
1.2.1 Terminology	2
1.2.2 Problems with traditional data processing systems that gave rise to the development and implementation of data base technology	4
1.2.3 Data base systems	5
1.2.3.1 Data base defined	5
1.2.3.2 How data base systems differ from traditional systems	6
1.2.4 Information resource management	7
1.2.5 The impact of corporate data base technology on control and systems development	8
1.3 Computer audit responsibility	9

1.4	Reasons for undertaking the study and research methodology	10
1.5	Problem description	11
1.6	Objectives of the study	11
1.7	Diagrammatic representation of the study	12
1.8	Hypothesis	12
1.9	Delimitations	12

## **CHAPTER 2 : INFORMATION RESOURCE MANAGEMENT**

2.1	Introduction	14
2.2	Overview of information resource management	15
2.2.1	Information resource management defined	16
2.2.2	Objectives of information resource management	17
2.3	Factors that led to an information resource management approach	18
2.3.1	The information age	20
2.3.2	Data processing growth and maturity	21
2.3.3	The need for management information	25

2.3.3.1	Anthony's management paradigm	25
2.3.3.2	Management information system concepts	29
2.4	An information resource management emphasis for organisations	31
2.4.1	Information as a resource	31
2.4.1.1	Information life cycle	34
2.4.1.2	Cost and value considerations	38
2.4.2	An information resource management emphasis	38
2.4.2.1	A proactive approach	39
2.4.2.2	Strategic information planning and the development of mission critical systems	39
2.4.2.3	Information technology architecture	43
2.4.2.4	Centralised information management	44
2.5	Summary	45

### **CHAPTER 3 :    SYSTEMS DEVELOPMENT WITHIN A CORPORATE DATA BASE ENVIRONMENT WITH SPECIFIC REFERENCE TO DATA RESOURCE MANAGEMENT**

3.1	Introduction	47
3.2	Data resource management	48
3.2.1	Data resource management philosophies	49
3.2.2	Data resource management defined	50

3.2.3	Data management function	51
3.2.3.1	Data administration versus data base administration	52
3.2.4	The data base environment	53
3.2.4.1	Overview of the data base approach	53
3.2.4.2	The components of a data base environment	57
3.3	Systems development in a data base environment	61
3.3.1	The role of systems development in general	61
3.3.2	The impact of a data resource management emphasis on systems development	62
3.3.2.1	A data-driven emphasis to systems development	63
3.3.2.2	A data base development life cycle	64
3.3.3	The interaction between the data base development life cycle and the traditional systems development life cycle phases	72
3.3.3.1	Requirements analysis and feasibility phase	73
3.3.3.2	Systems design phase	74
3.3.3.3	Program design and coding phase	75
3.3.3.4	Testing and acceptance phase	76
3.3.3.5	Operations and maintenance phase	76

3.4	Systems development responsibilities within the context of a formal data management approach	77
3.4.1	Data administration	79
3.4.1.1	Determining the information needs of the organisation	79
3.4.1.2	Creating and maintaining the corporate data models and ensuring that they are as stable as possible	80
3.4.1.3	Obtaining agreement among users about definitions and format of data	81
3.4.1.4	Ensuring that systems builders conform to the data models as far as possible	81
3.4.1.5	Resolving conflicts about incompatible representations of data	82
3.4.2	Data base administration (DBA)	82
3.4.2.1	The design of the data base	83
3.4.2.2	Maintaining data integrity, security, accuracy and completeness	83
3.4.2.3	Coordinating computer operations	84
3.4.2.4	Monitoring and improving system performance	84
3.4.2.5	Providing administrative support	85
3.5	Summary	85



**CHAPTER 4 : THE IMPACT OF A FORMAL DATA MANAGEMENT  
APPROACH WITHIN THE CONTEXT OF A CORPORATE  
DATA BASE ENVIRONMENT ON SYSTEMS DEVELOPMENT  
AUDITING AND AUDITING IN GENERAL**

4.1	Introduction	87
4.2	The role of the auditor in the evaluation of the system of internal controls during systems development	88
4.2.1	Internal controls defined	89
4.2.2	The external auditor's responsibility regarding internal control	92
4.2.2.1	Auditor responsibility for the evaluation of internal control according to general auditing standards issued by the South African Institute of Chartered Accountants	93
4.2.2.2	Auditor responsibility for the evaluation of internal control in a computer environment according to the audit and accounting guide issued by the South African Institute of Chartered Accountants	94
4.2.2.3	Auditor responsibility for the evaluation of internal control in terms of supplementary auditing standards relating to EDP environments issued by the American Institute of Certified Public Accountants	96

4.2.2.4 Auditor responsibility for the evaluation of internal control in terms of supplementary auditing guidelines relating to EDP environments issued by the International Federation of Accountants	99
4.2.3 The internal auditor's responsibility for internal control in manual and automated environments	101
4.2.4 Modern auditing practice pertaining to audit involvement during systems development	104
4.2.5 Internal/External auditor cooperation in participating in systems development	107
4.2.6 The auditor's objectives for participating in the systems development process	109
4.2.7 The scope of the auditor's participation in systems development process	110
4.3 Ensuring the reliability of an organisation's information systems with specific reference to data integrity and data management	111
4.3.1 Ensuring data integrity in a data base environment	111
4.3.1.1 Data integrity described	111
4.3.1.2 The importance of data integrity within a data base environment	113

4.3.2	Primary risk related to information system development within the context of an IRM approach to systems development	114
4.3.3	The role of data administration in ensuring the integrity of an organisation's data	116
4.4	The impact of a formal data management approach within the context of a data base environment on the organisational control structure and auditing	118
4.4.1	Control objectives and risks	118
4.4.2	Internal control	119
4.4.3	The impact of formal data management on the organisational internal control structures	121
4.4.4	The impact of formal data management on the auditing process in general	123
4.5	The impact of a formal data management approach in the context of a data base environment on the traditional systems development audit process	125
4.5.1	Data management and logical data integrity described in terms of the scope of this study	126

4.5.2	The impact of data management on the traditional systems development audit process per life cycle phase	129
4.5.2.1	Initiation phase	130
4.5.2.2	Analysis phase	131
4.5.2.3	Design phase	135
4.5.2.4	Construction phase	141
4.6	Summary	144

## **CHAPTER 5: CONCLUSION AND RECOMMENDATIONS**

5.1	Introduction	145
5.2	Research findings per identified objective	147
5.2.1	Objective 1: To ascertain the effect of corporate data base technology and data management practices on the traditional systems development process	147
5.2.1.1	Obtaining an understanding of data base technology and data management practices within an overall IRM approach to information management	148
5.2.1.2	Determining the effect of a data management approach on the traditional systems development process in the context of a data base environment	149

5.2.2	Objective 2: To determine the impact of data base technology and commensurate data management practices on the systems development life cycle audit process and systems auditing in general	150
5.3	Conclusions	152
5.4	Recommendations for future research	153
	<b>BIBLIOGRAPHY</b>	154
	<b>APPENDIX 1: DATA MODELLING CASE STUDY</b>	166
	<b>APPENDIX 2: GLOSSARY OF DATA TERMINOLOGY</b>	188

**LIST OF FIGURES**

<b>FIGURE NO.</b>	<b>DESCRIPTION</b>	<b>Page</b>
1.1	Structure of the study	
2.1	The six stages of data processing growth by Nolan	22
2.2	Anthony's management paradigm	27
2.3	The management process	33
2.4	Organisational functions supported by resource management functions	35
2.5	The resource life cycle	36
2.6	An information engineering approach to systems building	41
3.1	The collection of data in a data base	54
3.2	Various views of data in context of a data base environment	56
3.3	The components of a data base environment	58
3.4	The data base development life cycle	65
3.5	A data architecture framework	67
3.6	An entity-relationship model : Mountain View Community Hospital	69
3.7	Interaction between the data base development life cycle and systems development life cycle	73
4.1	Controls in an EDP environment	92
4.2	The control environment	120
4.3	Representation, descriptions and models supporting an information systems architecture	127
4.4	Components of the conceptual data model	128

**LIST OF TABLES**

<b>TABLE NO.</b>	<b>DESCRIPTION</b>	<b>Page</b>
2.1	Examples of activities in a business organisation classified according to Anthony	28
2.2	Examples of resources used by organisations	34
3.1	Responsibilities and interactions between organisational groups involved in systems development within the context of a formal data management approach	78

# **THE IMPACT OF THE DATA MANAGEMENT APPROACH ON INFORMATION SYSTEMS AUDITING**

## **CHAPTER 1**

### **OVERVIEW**

#### **1.1 INTRODUCTION**

Many knowledgeable data processing managers view data base technology as the single most important computer-related phenomenon since operating systems came into being. Those involved in the field of data base technology are constantly challenged by its burgeoning nature and the pace of new developments.

One of the latest developments in the field of data base technology is the recognition of information and its underlying data as a strategic resource and the commensurate management thereof – the concept known as information resource management.

Most of what has been published and spoken about information resource management to date has been prompted by the evolution in data base technology. The concepts of data base technology and information resource management are very intertwined. One must, however, remember that what has too often been left unsaid is that information resource management's reason for being actually falls within the ambit of the management discipline. The importance of information management lies at the very centre of the management process itself since the essence of management is rational decision-making based on the best available data. (Porter and Perry 1984:40).

The management of information has become a key process within every organisation. The areas of information resource management such as data resource management, which includes data and data base administration, all



present unique management issues, as well as audit and control considerations.

The implementation of information and data resource management principles in the organisation introduces a new consideration of the auditing of information systems that requires both a new approach to auditing as well as techniques.

## **1.2 THE CONCEPTS OF A CORPORATE DATA BASE ENVIRONMENT**

### **1.2.1 Terminology**

An entire vocabulary of terms describing the new technology and disciplines has developed with the emergence of managing data as a resource. These concepts are dealt with in Chapters 2 and 3 as well as the glossary list (Appendix 1). For the purpose of an introduction to the study, the following concepts are described:

- **Data**

Data can be defined as raw facts as a result of observation or measurement.

- **Data administration**

Data administration is in general concerned with the "ownership" of the data resource, its meaning, its relationship with other data and its global integrity. The data administration function is responsible for the high-level planning, modelling and coordinating of the organisation's data resource.

- **Data base**

A structured store of data containing data held in a structure that allows it to be retrieved using the contents of the data and its relationships.

- **Data base administration**

Data base administration is concerned both with the way in which the abstract data models are physically implemented in a computer system and with controlling the physical data base, that is the "custodianship" role over data.

- Data base management system

A set of software facilities to create, access and control a data base.

- Data base systems

Data base systems can be described as application systems, clerical procedures, application programs and data base(s).

- Data base technology

Data base technology can be described as the data management systems and disciplines that support an information and data resource management approach which pertains to an organisation's information resources.

- Data resource management

Data resource management can be defined as the function administering the organisation's data as a resource, and thus as an asset, in execution of the business goals. It includes both data administration and data base administration.

- Data management systems

Data management systems can be described as the software facilities for the management of data pertaining to data base systems. These facilities include data dictionaries, data base management systems and distributed data base management systems.

- Information

Information is data that have been processed into a meaningful form to the user.

- Information resource management

Information resource management can be defined as the function of managing the information resource (data processes, people and technology) of an organisation as a corporate resource in support of the organisation's goals.

### 1.2.2 Problems with traditional data processing systems that gave rise to the development and implementation of data base technology

When the computer was first introduced to business during the 1950's, basic business functions were merely automated. The entire effort took place with minimal planning and control and data processing rendered a useful service which assisted businesses in running more efficiently (Holloway 1988:5).

Roberts (1985:7.2) explains that as experience grew the number of applications increased very quickly. During the 1960's systems became partly integrated, with mutual data feeding. Roberts points out that with the interdependency of subsystems the enhancement and maintenance of systems became problematic and very costly. Systems also tended to become inflexible as regards changing management reporting requirements.

Roberts (1985:7.3) furthermore points out that these systems were designed to achieve specific circumscribed objectives, for example to produce certain programmed reports. They were designed application by application, with each application having its own master and transaction files. The systems were rigid and cumbersome, since many applications processed the same files.

The following representative summary of the problems encountered with the systems during the late 1960's and early 1970's is given by Roberts (1985:7.3):

- frequent sorting of files – it was not unusual for 25-30 per cent of processing time to be spent on the sorting of elements

- duplication of data elements in several files – this led to the inefficient use of storage facilities and difficulty in keeping data at the same degree of currency
- a proliferation of master files and transaction files throughout the organisation occurred, making it difficult to manage, maintain and store the large numbers of data files
- up-to-date management data was often not readily available

The search for solutions to these problems and the realisation that it was not the computer but rather the data stored in the computer that was the prime asset resulted in the idea and development of data base systems that support organisations on a corporate-wide basis.

### 1.2.3 Data base systems

#### 1.2.3.1 Data base defined

Martin (1976:4) defines a data base as:

*a collection of interrelated data stored together with controlled redundancy to serve one or more applications in an optimal fashion; the data are stored so that they are independent of programs which use the data; a common and controlled approach is used in adding new data and modifying and retrieving existing data within the data base.*

Techavichit (1979:5) explains that the term "data base" became popular in the early 1970's. Before that the term "application files" was virtually the only term used to refer to a collection of data. He adds that application files are collections of data associated with specific applications.

His conclusion is that a collection of various application files, which is often referred to as a "data base", does not mean the same as the "data base", which is used within a data base system.

The study will use Martin's definition of data base.

#### 1.2.3.2 How data base systems differ from traditional systems

According to Date (1977:4) the basic understanding of the data base system is that the data base is a (logically central) repository of all the data needed by an entire organisation, not just for one specific user or group of users.

The major changes brought about by the introduction of data base management systems have to do with the way in which information is organised, used and managed. It is not an evolutionary development but a complete revolution in information management and usage, as Roberts (1985:7.3) professes.

Traditional or conventional systems are application file systems specifically designed to satisfy specific output requirements, with the data input, storage and processing techniques geared to the specific output needs of the application program. For each new application a new file is created. Martin (1980:13) explains that in a large installation hundreds or even thousands of such files exist, with concomitant redundancy of data. When an application is changed and its file has to be restructured, all the programs which use that file have to be changed.

Within the ambit of data base systems analysis the business data is analysed, before being stored, in terms of basic elements and their inherent structure in order to establish data models that represent the inherent logic of the organisation's business and support the output requirements of various business functions and ultimately the organisation as a whole. Because many different users may require the same data stored in the data base, individual subsets or "views" of the data base may be shared by numerous users. This concept of "sharing" by multiple users therefore requires that the data be "independent" from any specific application (Techavichit 1979:1-2).

#### 1.2.4 Information resource management

Curtice (1986:99) maintains that although we all have our own ideas of what a data base is, the general concept behind developing a data base does not have a widely accepted definition. He points out that the data base concept is very much intertwined with the concept of information resource management.

Managing information as a resource is becoming more important on manager's agendas. Knight and Silk (1990:227) explain:

*The value of information itself, and the power of modern information technology for handling it, make this a more important subject for senior management than it was even five years ago.*

Information resource management is one of the most significant topics concerning information systems being discussed. It is being debated along various lines that include business system planning, information system analysis, design, development, data base design and quality assurance. Information resource management is an organisation-wide management programme that is corporate-oriented in its aim and scope and is directed at improving the quality of the organisation's data and information (McFadden & Hoffer 1991:6-8; Porter & Perry 1984:3).

Curtice (1986:99) argues that information resource management implies developing and managing organisational data bases that serve many different applications and users and has the ability to:

- share consistent and timely data throughout the business
- give users direct access to data
- evolve easily so that it meets changing business requirements

Information resource management is global in its objectives and scope, a key

feature being its centralised control of all the organisation's information resources (data, people, processes and technology)(Weldon 1990:3; Sibley 1983:2).

Since data is one of the primary resources that needs to be managed within an information resource management approach organisations are also starting to practice data resource management to ensure that data support the organisation's information objectives. Holloway (1988:1) explains that the key concept in this approach has been that data, like the other resources of the organisation, should be treated as a critical resource that any organisation has to manage.

The discussions of this study will address the discipline of data resource management within an overall information resource management approach to information system development.

#### **1.2.5 The impact of corporate data base technology on control and systems development**

Roberts (1985:7.4) explains that with conventional systems each functional application, such as payroll or accounts payable, is a separate subsystem with its own transaction and master files, reports and controls, the files being inflexibly bound to the program logic. The programmer includes whatever controls the subsystem will have in the application program. In the data base environment, however, many controls are centralised, the files are no longer tied to programs, and the information system is designed to best serve the needs of all information users. The programmer no longer "manages" the files the particular program uses - these functions are now handled by the data base management system software, which is under the control of an official data management function.

Roberts (1985:7.2) concludes that the concepts of data base technology essentially change the following two features of information systems:

- the methods of information control and usage, which change from an application-oriented approach to an enterprise-wide one
- the manager's perspective of information, which shifts from a functional reporting orientation to a corporate "decision-support-system"

Moss and Due (1988:12) emphasise that proper protection mechanisms must be established in the overall data base design to allow access to data resources and to ensure that control is maintained over integrity and auditability of these resources. Control requirements must be defined before storing data in a shared data base. Data base managed data can be protected when controls are designed to combine data base management system technological features with administrative procedures established for the organisation's general data processing environment and for its specific applications. A solution is only as good as the design. The data will be secured by the good design as well as proper overall data management practices. (Moss & Due 1986:2-3).

### 1.3 COMPUTER AUDIT RESPONSIBILITY

Generally Accepted Auditing Standard (AU230) entitled Accounting Systems and Internal Control, the South African audit and accounting guide named Auditing in a Computer Environment and its counterpart standards for the Institute of Internal Auditors, Standard 300, require that there be an appropriate organised study and evaluation of the internal controls on which the auditor subsequently relies in determining the nature, extent and timing of auditing procedures (Reilly 1986:8; SAICA 1986, SAICA 1989, IIA 1978:17).

In most advanced systems today, internal control procedures are built into the system and function with little or no human intervention.

Chen (1983:74) explains that in encountering these large, complex and even highly integrated systems the auditor finds it difficult to make a



knowledgeable and thorough evaluation of built-in controls after implementation or within the time typically available. He points out that a lengthy study of systems after implementation would most likely attract management resistance, and that it is therefore feasible for auditors to be "involved" in systems during their development phases.

In line with this thinking and modern auditing practice, Kothari (1988:55) emphasises that the most effective approach to evaluate systems is still to participate during the development process. He points out that there is a far greater chance of detecting any functional shortcomings or control weaknesses early enough with the auditor involved right from the start. It is easier to make changes to systems during their development than after they have become operational, because modification may not be practical or even viable, Gallegos and Dow (1987:1) stress.

In terms of the auditor's role in systems development there is no doubt that the data base management approach considerably impacts the audit approach to systems development since the introduction of new data structures, access concepts, control techniques and system software components now affect the processing of significant (accounting) transactions. The auditor must resist the temptation of insisting on traditional controls that could nullify many of the efficiencies to be gained from this corporate data base approach. The auditor must carefully examine the new emphasis, techniques and approaches and ensure that the proper controls are provided.

#### **1.4 REASONS FOR UNDERTAKING THE STUDY AND RESEARCH METHODOLOGY**

Although recent literature on data processing, information systems, computer science and auditing includes many data base technology references, very little research into the impact of data management practices on systems development and the auditing thereof has been published. The methodology adopted for this study is a survey of authoritative literature, in respect of each of the mentioned sub-problems in Section 1.5 that can have

a direct or indirect impact on offering a solution to the sub-problems, and a substantiation of the hypothesis. Through an analysis of the findings and facts gathered, a conclusion is reached on each sub-problem.

This type of research will ably serve the purpose of providing auditors of all descriptions, both internal and external, knowledge of the impact of modern data management practices on systems and systems development auditing.

### **1.5 PROBLEM DESCRIPTION**

Changing technology has always posed problems for auditors, and data base technology is no exception. The challenge to auditors is to be in control of the technology and to ensure that systems are built to take advantage of the potential benefits of technology while also improving the traditional internal controls to make systems reliable and secure.

As organisations are increasingly starting to utilise data bases and older applications are being converted to take advantage of the up-to-date data base technology, it is becoming important for auditors involved in computer audits to:

- identify and define the effect that corporate data base technology and sound data management practices have on the traditional systems development process (sub-problem 1)
- determine the impact that proper data management practices have on the auditing of systems (sub-problem 2)

### **1.6 OBJECTIVES OF THE STUDY**

The objectives of this study are to:

- ascertain the effect that corporate data base technology and data management practices have on the traditional systems development

process in the context of an information resource management approach to systems building. (These aspects of the study will be dealt with in Chapters 2 & 3)

- determine the impact that data base technology and data management practices have on the systems development life cycle audit and systems auditing in general. (This will be discussed in Chapter 4.)

## **1.7      DIAGRAMMATIC REPRESENTATION OF THE STUDY**

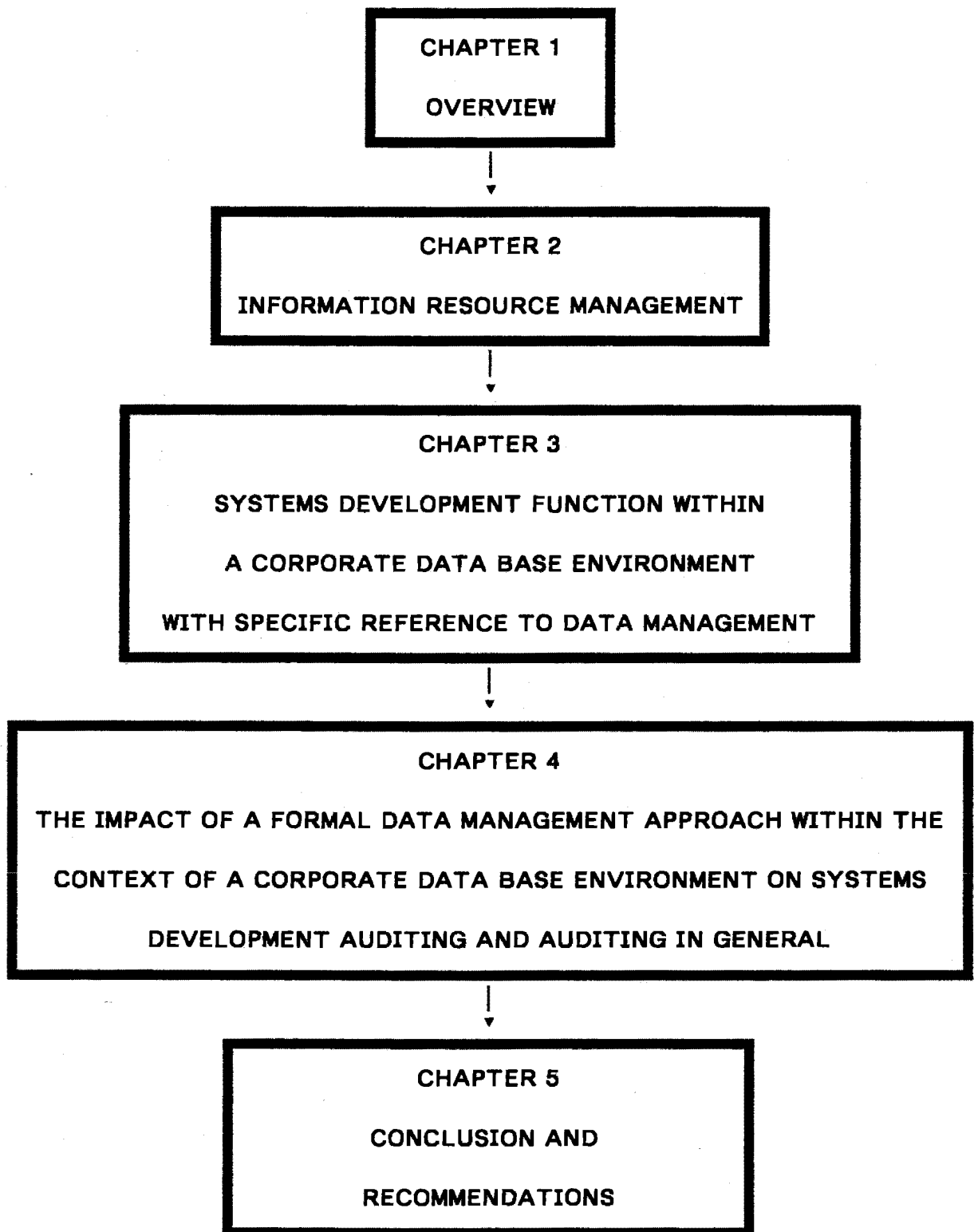
A representation of the study is given in Figure 1.1.

## **1.8      HYPOTHESIS**

The use of corporate data base technology and data management practices introduces a new consideration of the systems development audit process in that the auditor must adopt a more holistic approach to the auditing of information systems. The emphasis is to view the audit of any specific system as a subsystem (and not as a stand alone system) of a bigger system that utilises and updates a corporate data base in its support of the mission of the organisation.

## **1.9      DELIMITATIONS**

The study concentrates on the impact of formal data management practices within the context of a corporate data base environment on systems and systems development (compliance) auditing. Substantive auditing is not covered. The discussions take place within the framework of an information engineering approach to systems development. The discussions will focus on the logical aspects of data base technology and data management. The study does not address the technical aspects of data base technology as well as the commensurate physical integrity aspects and controls of data.

**FIGURE 1.1: STRUCTURE OF THE STUDY**

## **CHAPTER 2**

### **INFORMATION RESOURCE MANAGEMENT**

#### **2.1 INTRODUCTION**

It is impossible to discuss the concepts of data management and corporate data base technology and their impact on auditing, which is the premise of this study, without first addressing the concept of information resource management (IRM). IRM is the guiding principle behind the development of data bases, the implementation of corporate data base technology and data management practices.

Curtice (1986:99) explains that the notion of a data base is a familiar one in the modern information systems practice. He argues that although most of us have our own ideas of what a data base is, the general concept for developing a data base - apart from particular applications - has no widely accepted definition. He points out that a data base as a guiding principle for developing specific information systems, is very much intertwined with the concept of IRM - treating information (and its underlying data) as a corporate resource.

Data processing has evolved over the years and the information era in which organisations are reaching towards maturity has now been entered. Today, systems are being developed to assist management in making decisions that have significant importance for the ability of the organisation to run profitably, compete, or even survive.

As organisations become more able to cope with the growing pressures of modern business, the providers of information have an increasingly important role to play in supporting the organisations in the achievement of their goals. Organisations are only just beginning to recognise the value of IRM.

The connection of data base technology implementation with IRM implies developing and managing corporate data bases that serve many different applications and users with consistent and timely data in support of the organisation's information requirements (Curtice 1986:99). Appleton (1983:225) explains that in this regard the aim of IRM is to manage, store, give access to, and provide the ability to manipulate and to communicate the raw material of information : data.

In this chapter it is proposed to:

- give a short overview of the IRM concept
- discuss the factors that gave rise to an information resource management approach, namely
  - the information age
  - the evolution in data processing – from an emphasis on the hardware to a data and information resource management emphasis
  - the need for better management information
- discuss the guiding principles for an IRM emphasis for organisations

## **2.2 OVERVIEW OF INFORMATION RESOURCE MANAGEMENT**

Over the past several years management has been trying to figure out how best to manage the corporate information resources. It has generally still not realised the value of explicitly managing information, which is the essence of an idea which is at least 10 years old and is appropriately called information resource management (IRM) (Bryce 1987:89).

The guiding principle of IRM is the concept of asset management –treating information as a corporate resource. Its primary aim is to meet, with

increasingly effective results, the needs of the enterprise for information (Porter & Perry 1984:39; Curtice 1986:99).

IRM is the management of the data, people and processes that produce information serving a business or functional need. Data is a key component of IRM and is of strategic importance to an organisation. For this reason organisations are starting to practice data management to ensure that data is available and structured to support corporate objectives (Bryce 1987:89; Meador 1988:9).

In order to acquire a broad understanding of the IRM concept, it is proposed to:

- define the IRM approach
- briefly discuss the objectives of IRM

#### **2.2.1 Information resource management defined**

There seems to be little consensus in information and computer management literature about the definition of IRM.

Some authorities view it simply as the management of information as one of the fundamental resources of the organisation such as personnel, equipment and capital (Durell 1985:1). There are also those that believe that it is the management of the resources of information, that is data, people, processes, and its underlying technology (Bryce 1987:89).

One can argue that these viewpoints mean the same. Looking at the human resource management function more closely one finds that this function also implies that the underlying "resources" of human management such as managerial talent and labour also need to be managed (Griffin 1987:8). The conclusion is therefore that the management of information implies management of its underlying resources as well.

Porter and Perry (1984:40) define IRM as follows:

*Information Resource Management (IRM) is whatever policy, action, or procedure concerning information (both automated and nonautomated) that management establishes to serve the overall current and future needs of the enterprise. Such policies, etc., would include considerations of availability, timeliness, accuracy, integrity, privacy, security, auditability, ownership, use, and cost effectiveness.*

In line with Porter and Perry, the British Computer Society's Data Base Specialist Group (1989:2) defines IRM as the approach or policy which an organisation adopts to manage its data / information resource.

The British Computer Society's definition covers:

- the actual formal information an organisation chooses to account as its information resource
- the information systems designed to provide the information to the various groups and individuals who request and/or need the information
- the information technology used to manage the information resource
- the implementation approach to introduce both information systems and information technology in the organisation

In terms of this study, based on the above discussion, IRM will be defined as the function of managing the information resources of an organisation as a corporate resource in support of the organisation's goals.

### **2.2.2 Objectives of information resource management**

IRM is an organisation-wide management programme that is global in its objectives and scope. Like all management programmes that span over the whole organisation, the IRM programme is holistic in its strive for information



consistency, completeness and compatibility. (Porter & Perry 1984:40)  
Again it can be compared to the Human Resource Management function that is also global in its objectives of creating synergy in the organisation in terms of the management of its people.

Porter and Perry (1984:39) claim that the primary objective of IRM is to ensure that the organisation's information needs are met with increasingly effective results.

Bryce (1987:89) explains that it means information systems that produce meaningful and timely information to the users who can utilise the information in an intelligent way to carry out the organisation's mission, objectives and responsibilities in a cost-effective manner are in place.

### **2.3 FACTORS THAT LED TO AN INFORMATION RESOURCE MANAGEMENT APPROACH**

Since the first computer was installed in business in 1954, the growth has been phenomenal. Computers have brought about irreversible change in management, operations and structure of business (Dawley 1983:15). The 1980's heralded the information revolution, which was necessitated by the increased importance of information in the Western economy (Synnott & Gruber 1981:75).

The change from a manufacturing society to an information based society has been accelerated by information technology developments. The changes have been evolutionary and have developed into what is today known as the information era. The previous hardware and software emphasis of the computer era is shifting toward a focus on information management (Synnott & Gruber 1981:3; Holloway 1988:2).

Synnott and Gruber (1981:3), explain:

*This difference in focus is significant because it emphasizes the quality and the value of the output of computers rather than the*

*quantity, "by-the-pound" approach of the past. The data processing industry has demonstrated its capacity to produce huge outputs of paper – literally billions of pages every working day. What is needed, however, is not more paper or raw data, but distilled, summarized information that can be accessed, assimilated, and used more effectively, particularly by managers.*

In discussing this new "information resource" management emphasis it is important to understand the factors and developments that have taken place in the world economy, the computer industry and in terms of the information management practices that gave rise to the IRM approach.

Naisbitt (1984:1) explains that as a society we have been moving from the old to the new and are still in motion. Although we may continue to think we live in an industrial society, we have in fact changed to an economy based on the creation and distribution of information.

- The information age concept will be discussed under Section 2.3.1.

There has also been a shift in emphasis that has taken place in the data processing industry itself. Nolan (1979:117) points out that organisations can only obtain data processing maturity by adopting a resource management approach.

- Nolan's stage model will be discussed under Section 2.3.2.

When addressing the factors that led to an information resource management approach they cannot be separated from the growing needs of management for information. It is therefore proposed to:

- discuss the three-level paradigm of Anthony since his model still forms the basis for determining management information requirements. (Knight & Silk 1990:11; Martin 1976:301; Holloway 1988:4 and Duffy & Assad 1980:12). Section 2.3.3.1 will address Anthony's paradigm

- discuss the management information system (MIS) concepts. Section 2.3.3.2 will give an overview of management theory and decision making. Zani's assertion will be introduced as a way of solving management's information systems requirements

### 2.3.1 The information age

Naisbitt (1984:49) refers to the fact that approximately 200 years ago we were basically an agricultural society involved in the growing of food to feed ourselves. At the turn of the nineteenth century we evolved into an industrial society where the success of an organisation was measured by how many goods and products could be produced by an assembly line or in a factory. We were involved in the industrial society until the 1950's, when we entered the computer era (Holloway 1988:2).

With the advanced technology that is available today, we are experiencing the continuing evolution and integration of technologies such as networks, workstations, software and hardware into an integrated philosophy collectively termed information technology (Morton 1991:4).

Knight and Silk (1990:139) acknowledge the fact that by the end of the 19th century improvements in communications had reduced the previous isolation of countries. Telegraph cables spanned the developed world, although the telephone was still in its infancy.

Information technology, supported by advances in computers, software systems and communication networks has now made it possible for global organisations to operate 24 hours a day, seven days a week, across immense geographical distances. In 1986 the Society of Worldwide Interbank Financial Telecommunications (SWIFT) reported that 800 000 messages were passing through their global banking network every 24 hours (Etheridge 1986:64).

Schatz (1988:41) reports that only recently the executive office of the president of the USA announced that a further technology revolution was on

its way – their management letter predicts:

*By the year 2000, almost two thirds of American households are expected to own personal computers, and 6-8 million businesses and 40-50 million households will have electronic access to databases containing information on available products and services from private and public organisations.*

Businesses are today realising that they can utilise information produced by computers to achieve corporate success and growth. The industry has now evolved into what is known as the information age (Holloway 1988:2).

### **2.3.2 Data processing growth and maturity**

The role of data processing (DP) in an organisation has attracted much attention over the years. Nolan's stage model (1979:117) is perhaps the most well known and widely cited model of the DP evolution in organisations. Nolan proposes six stages of DP growth from the original introduction of DP into the organisation (stage 1) to the mature management of data resources (stage 6).


The logic behind Nolan's model holds that the major activities in the management of computing are identifiable in stages that correspond with periods of stability along the growth path of data processing.

Nolan's paper identifies (1979:117) the following six stages in the data processing growth – (depicted in Figure 2.1 on the following page):

#### **STAGE ONE (INITIATION)**

This stage is the start of the growth process. Applications are primarily functional cost reduction applications and the users are typified as being "hands-off" (Nolan 1979:117; Synnott & Gruber 1981:7).

**FIGURE 2.1 THE SIX STAGES OF DATA PROCESSING GROWTH BY NOLAN**

Applications portfolio	Functional cost reduction applications	Proliferation	Upgrade documentation and restructuring of existing applications	Retrofitting existing applications using data base technology	Organisation Integration of applications	Application integration "mirroring" information flows
DP organisation	Specialisation for technological learning	User-oriented programmers	Middle management	Establish computer utility and user account teams	Data administration	Data resource Management
DP planning and control	Lax	More lax	Formalised planning and control	Tailored planning and control systems	Shared data and common systems	Data resource strategic planning
User awareness	"Hands off"	Superficially enthusiastic	Arbitrarily held accountable	Accountability learning	Effectively accountable	Acceptance of joint user and data processing accountability
Level of DP expenditures						
	<b>Stage I</b> Initiation	<b>Stage II</b> Contagion	<b>Stage III</b> Control	<b>Stage IV</b> Integration	<b>Stage V</b> Data administration	<b>Stage VI</b> Maturity

Source: Nolan, R.L. 1979. Managing the Crisis in Data Processing. Harvard Business Review. March-April. (page 117).

During this stage the basic business functions are automated mainly to reduce costs and to make organisations more effective. Data processing delivers a useful service that assists in running businesses more efficiently but the whole effort takes place with the minimum amount of planning and control. (Nolan 1979:117,118).

## STAGE TWO (CONTAGION)

This stage sees the contagion of the technology. A proliferation of applications takes place and the users become superficially enthusiastic (Nolan 1979:117). Most systems are expanded. There are systems and data all over the business, and each system has a unique set of data to be processed. The entire data processing and control effort during this phase is even more lax than during the previous phase. Much of the information referenced by these systems and programs is redundant information and it becomes very difficult to maintain currency. Management begins to realise that they have to take stock of the systems and information that is available (Holloway 1988:5; Nolan 1979:117).

## STAGE THREE (CONTROL)

At stage 3 formalised planning and control is introduced. Applications are upgraded, documented and restructured. Users are arbitrarily held accountable for their use of technology (Nolan 1979:117). Existing applications are restructured to try to make use of the information that already exists and information is passed from one system to the next.

Holloway (1988:6) explains:

*...in order to accomplish this restructuring, some middle-management control systems had to be introduced. But something more important happened in stage 3. Until this point, the focus had been on the computer, the computer was the asset of the organisation - the prime resource - the reason for being able to generate a lot of information for people. In getting into stage 3 and performing this restructuring operation, it was realized that it was not the computer which was the prime asset. It was the data stored on the computer, and this data could be very valuable to end users in accomplishing the running of the organisation.*

Once it is realised that data rather than the computer is the important asset, it becomes necessary to find ways to make the data available to the various systems throughout the organisation. As stage 3 draws to a close, a move

is made to data base and data communication technology. The move highlights a change in emphasis. The transition point where the focus shifts from the management of technology to management of data is depicted in Figure 2.5 (Holloway 1988:6; Nolan 1979:117).

#### STAGE FOUR (INTEGRATION)

Stage 4 sees existing applications being retrofitted using data base technology and users becoming more accountable for their applications (Nolan 1979:117).

Holloway (1988:6) elaborates:

*The goal here was to have one place in which to store corporate data. Instead of having multiple diverse architecture systems which required programming to access stored data, the information was stored in the database so that it could be readily available to everyone who needed it. In order to do this, some formal planning and control systems had to be instituted, and the end user was increasingly involved with helping with the information design and data definition.*

#### STAGE FIVE (DATA ADMINISTRATION)

Stage 5 sees the introduction of data administration and the further integration of applications which are typified by shared data and common systems (Nolan 1979:119).

All systems share a nucleus of data. Instead of having multiple technologies, architectures and multiple files which contain the same information, there is now one store medium for shared data to be used by all systems. Data administration is introduced to determine what information is to be captured, how it will be stored, who will have access to it and to assign security (Holloway 1988:6). By this stage, the users are effectively accountable.

## STAGE SIX

Stage 6, the final stage in Nolan's growth model, sees the maturity of data processing in the organisation. The applications portfolio is complete and its structure "mirrors" the organisation and the information flows in the company. Complete data resource management practices are now in place (Nolan 1979:117).

There is a strong drive to integrate the applications and to make information more readily accessible through technology during this stage as well as stages 4 and 5.

Nolan (1979:126) argues that stage 6 must be the aim of any organisation. To achieve this the data has to be managed and controlled as a resource.

Mach (1990:50) claims that the larger organisations have all introduced computers by now. Most major organisations have evolved to a point from either late in the control stage (Nolan's stage 3) to somewhere in the integration stage (Nolan's stage 4). Smaller organisations have not progressed this far. He argues that it is doubtful that any organisation, except perhaps some in Japan, has reached full maturity (stage 6).

Organisations have to adopt IRM principles or refine their information resource management principles to obtain maturity, Synnott and Gruber (1981:8) conclude.

### **2.3.3 The need for management information**

#### **2.3.3.1 Anthony's management paradigm**

During the early years of computerisation, that is the 1950's and early 1960's, there was no real management framework for computerisation. Organisations merely worked towards computerising routine administrative tasks. Typical applications were payroll, ledgers and inventory. These



were all developed on a historical bookkeeping basis with little or no "management" utilisation other than for the purposes of maintaining statutory records (Roberts 1985:7.2).

By the end of the 1960's a theory of business computing or electronic data processing (EDP) had begun to emerge. During this period the conceptual management framework in which much of the systems development and discussion took place was based on ideas derived from the paradigm, or model of business that Anthony (1965:19) developed.

According to Anthony (1965:22), the enterprise may be seen in terms of a trinity consisting of three main processes of strategic planning, management control and operational control.

These three processes are displayed in terms of a triangle as depicted in Figure 2.2.

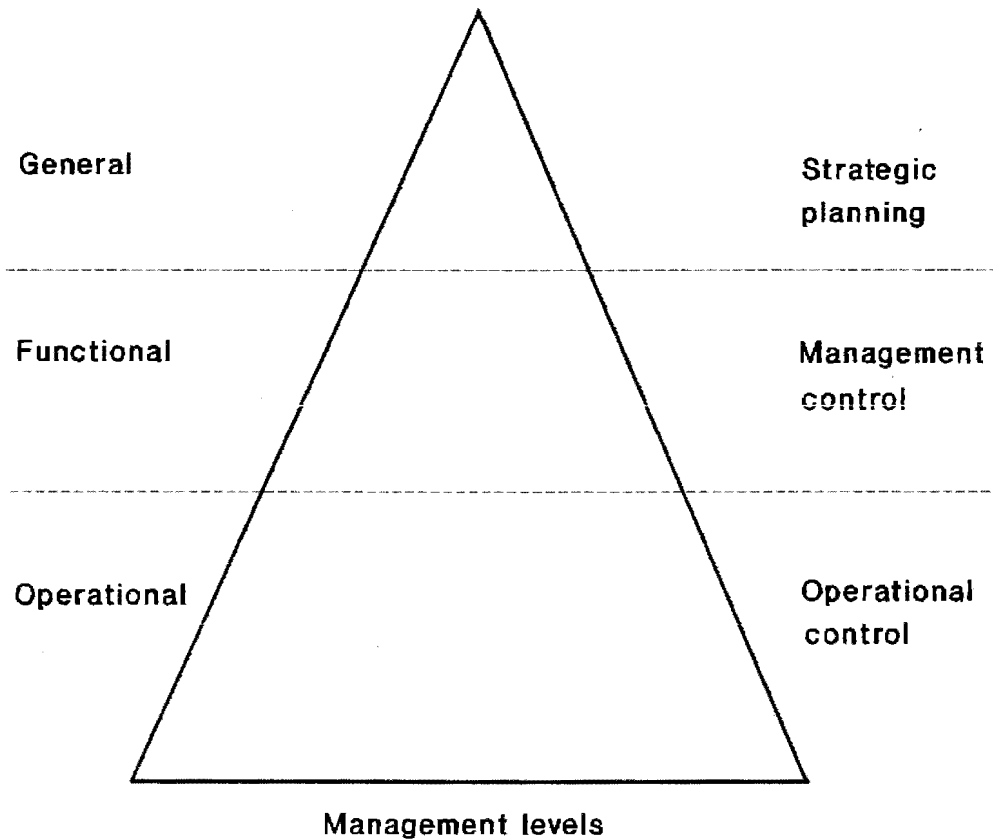
In this paradigm of Anthony (1965:15-17) the following three processes are defined:

#### STRATEGIC PLANNING

This is the process of setting the organisation's goals, of deciding upon the level of resources required and of determining the policy by which these resources will be used.

#### MANAGEMENT CONTROL

This is the process by which the organisation assures that the resources (identified during the strategic phase) will be efficiently and effectively employed.

**FIGURE 2.2 ANTHONY'S MANAGEMENT PARADIGM**

Source: Holloway, S. 1988. Data Administration. Vermont (USA): Gower Publishing Company. (page 4).

### OPERATIONAL CONTROL

This is the process used by an organisation to ensure that the detail of the work to be performed is efficiently monitored.

Anthony (1965:20) emphasises that organisations require planning and control systems to facilitate the three processes and to ensure that the organisations are kept on course. Computer systems are excellent tools for helping management to ensure that these processes are being performed.

The value of Anthony's classification is that it provides a practical terms of reference for analysing decisions, that is their level, their nature and who should be responsible. This provides a basis for identifying the information requirement's nature for each of these levels and a mechanism for evaluating existing information systems and structuring new ones (Duffy & Assad 1980:12).

Table 2.1 depicts certain examples of activities in a business as classified by Anthony.

**TABLE 2.1 EXAMPLES OF ACTIVITIES IN A BUSINESS ORGANISATION  
CLASSIFIED ACCORDING TO ANTHONY**

Strategic Planning	Management Control	Operational Control
Choosing company objectives	Formulating budgets	Controlling hiring
Planning the organisation structure	Planning staff levels	
Setting personnel policies	Formulating personnel practices	Implementing policies
Setting financial policies	Working capital planning	Controlling credit extension
Setting marketing policies	Formulating advertising programs	Controlling placement of advertisements
Setting research policies	Deciding on research projects	
Choosing new product lines	Choosing product improvements	
Acquiring a new division	Deciding on plant re-arrangement	Scheduling production
Deciding on non-routine capital expenditures	Deciding on routine capital expenditures	
	Formulating decision rules for operational control	Controlling inventory
	Measuring, appraising and improving management performance	Measuring, appraising and improving worker's efficiency

Source: Anthony, R.N. 1965. Planning and control systems: A framework for analysis. Division of Research, Graduate School of Business Administration. Boston:Harvard University Press. (page 19).

The basic premise of Anthony's framework is that the planning and control processes are focused on the internal efficiency of the organisation, for example the control of strategic planning involves the need for top management to monitor adherence to their stated strategies.

Duffy and Assad (1980:12) stress that the matching of an organisation's activities with Anthony's processes is not always clear-cut, but Anthony's approach has nevertheless proved to be useful.

#### 2.3.3.2 Management information system concepts

### MANAGEMENT THEORY AND DECISION-MAKING

There is no widespread agreement on the subject of organisational goals or objectives. Organisational goals have been stated to be the maximisation of profits, survival and even the current results of negotiations by the most powerful coalitions in an organisation (Duffy & Assad 1980:9).

Organisational goals are achieved by means of a process of resource conversion. In the process, inputs of people, finances, machines and materials and information are converted to outputs of products and services. In this process control is exercised by management (Duffy & Assad 1980:9).

Synnott & Gruber (1981:v) points out that information technology impacts on management's ability to potentially increase their effectiveness and productivity in managing the business. Griffin (1987:9) emphasises that, more specifically, information technology impacts on the traditional management functions of planning, organising, controlling and leading.

In this process of management a fundamental task of management is decision-making. Regardless of one's views on this matter, it is clear that decision-making is a fundamental link in a chain leading to the attainment of the organisational goals (Duffy & Assad 1980:9).

It is further clear that the information required by management must be such that it aids decision-making and helps managers execute their responsibilities effectively (Wysong 1983:30). This brings us to the discussion of management information systems.

## MANAGEMENT INFORMATION SYSTEMS

Zani (1970:98) developed a framework for producing management information systems (MIS) based on Anthony's paradigm that supports the functions of an organisation. Zani views MIS as being used to support strategic planning, management control or operational control.

Zani's application (1970:98) of Anthony's paradigm highlights the failure of data processing to meet the information needs of organisations. This led to the recognition of two types of information systems:

- those automating operational and control processes (operational control)
- those supplying decision-making information to managers (management control)

The corporate strategy should determine the organisation's information systems design objectives.

Zani (1970:98) asserts that this meant a new approach to information systems design, one that shifted the emphasis from operational systems toward systems that focused on the critical tasks and decisions made by managers.

In this regard Wysong (1983:30) points out that MIS are required both for day-to-day operations, measurement of performance and for tracking progress toward long-term strategic goals.

Wysong (1983:30) adds that MIS should provide top management with the following specific information that, which:

- is necessary to exercise control over operations
- will warn of developing problems early and indicate remedial action required
- will enable correct interpretation of financial statistics
- will enable effective allocation of human, financial, and mechanical resources

He concludes (1983:30):

*The organisations that are the most successful are those that recognize information as a major resource and structure it as efficiently as they do other assets. These organisations use information to make changes in plans in order to maximize cost-effectiveness or resources. Information is used to measure performance and responsiveness as an adjunct to profitability measurements. Information must be considered as a vital component of the organisation.*

## **2.4 AN INFORMATION RESOURCE MANAGEMENT EMPHASIS FOR ORGANISATIONS**

The previous sections of this chapter gave an overview of the historical development of business systems and the IRM concept. The discussions in terms of this section will now concentrate on assessing the role and importance of information as a resource in the context of the modern organisation.

### **2.4.1 Information as a resource**

Knight and Silk (1990:3) stress that managers today must run efficient, profitable and cost effective organisations. The assets that they have to manage to obtain results are the following:

- people
- money
- materials
- energy
- information

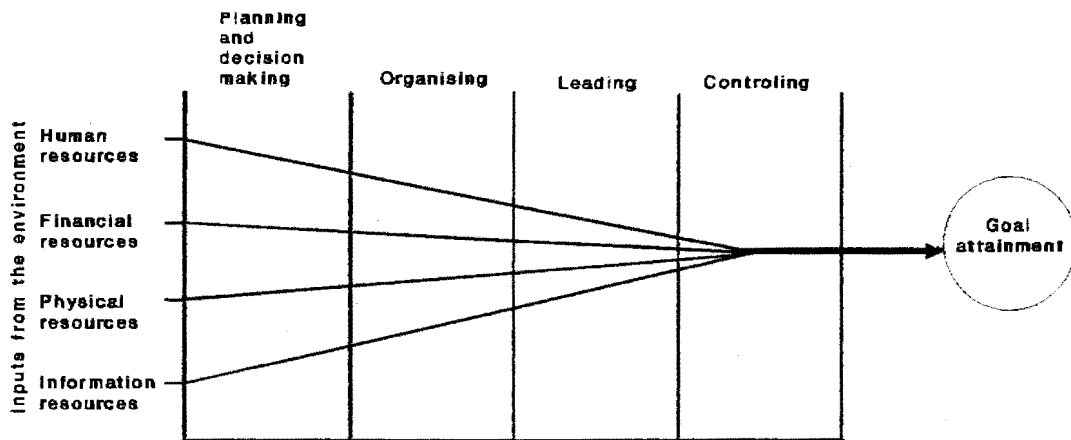
They claim that information is becoming increasingly important as a resource, for both the economy as a whole and individual organisations.

It has become important for businesses to manage their information resources with the same vigour as their more tangible assets, such as personnel, inventory and equipment, claims Durell (1985:1). Businesses are just beginning to feel the need for proper information asset management which is being driven by dramatic reductions in productivity, accompanied by even more dramatic increases in the volume and complexity of demand (Appleton 1986:71).

Information needs to be treated like the other assets of the organisation or corporation. It must furthermore be secured and controlled like cash. It has to be nurtured and utilised as effectively as the organisation's human resources. It also has to be accurate, Percy (1986:123) claims.

IRM objectives should follow the corporate objectives and goal attainment, Synnott and Gruber (1981:29) argue. This requires identification, review and continual updating of needs and development of information systems to meet those requirements with either internally generated or externally available data. It also requires a framework for the organisation and thus necessitates the formation of plans and policies to ensure consistency of the system with the organisation's mission, objectives, plans, priorities, legal constraints and security (Porter & Perry 1984:39).

FIGURE 2.3 THE MANAGEMENT PROCESS



Source: Griffin, R.W. 1987. Management. Boston:Houghton Mifflin. (page 9).

Griffin (1987:9) in his book Management describes that the objectives of resource management can be obtained through a process of planning and decision-making, organising, leading and controlling. This process is diagrammatically depicted in Figure 2.3.

Griffin (1987:8) explains that it is management's responsibility to combine and co-ordinate these various resources to achieve the organisation's goals. As an example he mentions a manager at Mobil Corporation, who utilises the talents of executives and platform workers, profits earmarked for reinvestment, existing refineries and sales forecasts to make decisions pertaining to the amount of oil to be refined for a specific period. Similar examples of resources used by organisations are shown in Table 2.2.



**TABLE 2.2 EXAMPLES OF RESOURCES USED BY ORGANISATIONS**

<b>Organisation</b>	<b>Human resources</b>	<b>Financial resources</b>	<b>Physical resources</b>	<b>Information resources</b>
Mobil Corp.	Drilling platform workers Corporate executives	Profits Stockholder investments	Refineries Office buildings	Sales forecast OPEC proclamations
University of Michigan	Faculty Secretarial staff	Alumni contributions Government grants	Computers Campus facilities	Research reports Government publications
City of New York	Police officers Municipal employees	Tax revenue Government grants	Sanitation equipment Municipal buildings	Economic Forecasts Crime statistics
Joe's Corner Grocery Store	Grocery clerks Bookkeeper	Profits Owner investment	Building Display shelving	Price lists from suppliers Newspaper ads for competitors

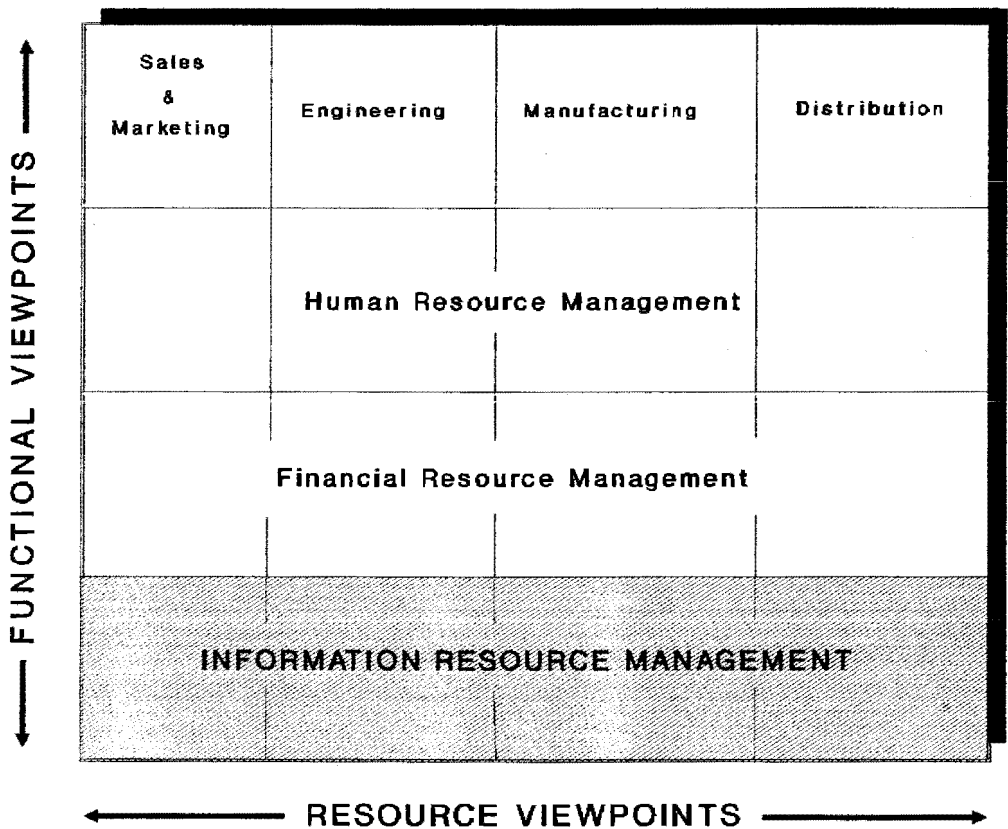
Source: Griffin, R.W. 1987. Management. Boston:Houghton Mifflin. (page 8).

The fact that organisations are recognising information as a resource signals a change in the historical role of data processing in that managers of technology are now becoming managers of data and information. It also signals the beginning of the final incorporation of data processing into the mainstream of business. Mach (1990:47) claims that IRM is starting to take its place as a proper resource management function within the organisation as is shown in Figure 2.4.

#### 2.4.1.1 Information life cycle

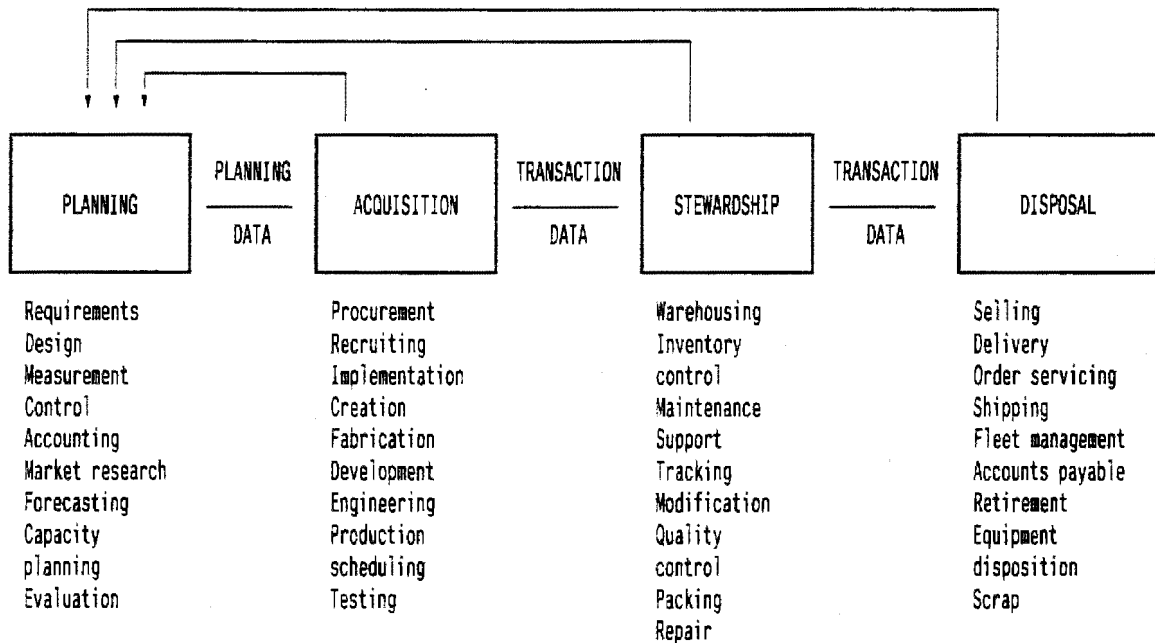
Information, like the other resources of the organisation, also has a specific life cycle (Meador 1988:9).

**FIGURE 2.4 ORGANISATIONAL FUNCTIONS SUPPORTED BY RESOURCE MANAGEMENT FUNCTIONS**



Source: Mach, R. 1990. Information engineers wanted to build enterprise models. Software Magazine. September. (page 44).

Martin (1990:56) explains that the functions that apply to products, services, or resources tend to be in the sequence of a four-phase life cycle, namely planning, acquisition, stewardship, and disposal, as is illustrated in Figure 2.5. Appleton (1983:25) points out that this life cycle is different from the systems life cycle that must be clearly understood and managed.

**FIGURE 2.5 THE RESOURCE LIFE CYCLE**

Source: Martin, J., 1990. Information engineering : Planning and analysis - Book II. New Jersey:Englewood Cliffs, Prentice-Hall, Inc. (page 56).

These phases can be explained as follows:

### PLANNING

Planning entails identifying the resource early enough so that it will be available when required Meltzer (1981:76).

Meltzer (1981:76) explains that although information is invisible, elusive and changeable, one can and must undertake long range plans for it. Brancheau & Wetherbe (1987:25) point out that information planning has been a topic of growing importance and ranks highly as one of the key issues for information technology (IT) management. This is mainly due to the information

revolution that is taking place and its impact in terms of:

- the rapidly changing business environment
- the accelerated technological change
- the application of IT for competitive edge

## ACQUISITION

Acquisition is the process of obtaining information for the organisation from either internal or external sources (Meltzer 1981:44).

Internal information varies in subject matter from accounting data to personnel records, based on all the organisation's data stored on the various data bases of the organisation. It consists of automated and non-automated information irrespective of the medium it is stored on (Meltzer 1981:44).

External sources include textbooks, handbooks, periodicals, reports, vendor catalogues, government sources and many others (Meltzer 1981:45).

## STEWARDSHIP (EMPLOYMENT AND MAINTENANCE)

This is the process of making information available through application systems end-user reports, and ad-hoc queries and maintaining the quality of information and its underlying data (Meador 1988:10).

## DISPOSAL OR DISPOSITION

This can be described as the aging, archival and retention process (Meador 1988:9).

#### 2.4.1.2 Cost and value considerations

Information is neither free nor cheap. Even the so-called "free" information obtainable from government sources costs organisations through paying taxes, Meltzer (1981:48) argues.

Information has a cost and a value. Knight and Silk (1990:6) refer to the following two approaches for estimating the value of information:

- the organisation's "willingness to pay" for information which is measured by the budget for books, periodicals and other information sources, plus the investment in time and effort to identify, retrieve and read the information
- to estimate the savings in cost or other advantages that result from the individual having the information

Meltzer (1981:48) emphasises that it is not easy to assign "actual" cost to information that is generated through a in-house system, but the fact remains that it costs money to produce information and therefore consideration should be given to assigning a value to information.

#### 2.4.2 An information resource management emphasis

The purpose of management is to ensure that an organisations' goals are attained in an efficient and effective manner, as Griffin (1987:8) explains:

*Management is the process of planning and decision making, organizing, leading, and controlling an organization's human, financial, physical, and information resources to achieve organizational goals in an efficient and effective manner.*

He adds that efficiency means doing things in a systematic fashion without unnecessary waste, while effectiveness means doing the right things

(1987:9).

The challenge of information resource management is to deliver the right information to the organisation in an efficient manner (Griffin 1987:8,9).

Martin (1989:1) claims that it means getting the right information to the right people at the right time. This new emphasis entails the following:

- a proactive approach to information management
- strategic information planning and the development of mission critical systems incorporating an information-engineering type of approach
- the necessity of an information technology architecture (ITA)
- centralised information management

#### 2.4.2.1 A proactive approach

To be able to deliver the right information to users requires a proactive approach (Synnott & Gruber 1981:34) claim. They explain that this can be achieved through proper systems planning in "searching out" opportunities to apply technology in solving business problems. Other writers claim that the main foundation for this approach is that the proactive use of information will create a competitive advantage (Borthick 1986:40,41; Morton 1991:82).

This proactive orientation is of course interpreted as one of the prerequisites for success within an organisation and a basis for action (Peters & Waterman 1982:121).

#### 2.4.2.2 Strategic information planning and the development of mission critical systems

Strategic information systems planning has its roots in traditional strategic

planning and places strong emphasis on how organisations can identify information systems that will give the organisation a competitive edge (Moeller 1988:3; Wiseman 1988:99).

Strategic information systems planning is the process of establishing a long-term "framework" for developing information systems in a way that they will optimise the effectiveness of an organisation's information resources and to enable them to support the objectives of the organisation (Bainey 1986:1).

Zani (1970:98) believes that the corporate strategy should determine the organisation's information systems design objectives. He explains that there should be a shift in emphasis from the operational systems towards systems that focus on the critical tasks and decisions made by managers within the organisation.

It follows that a proper resource management approach to information systems and strategic information planning should therefore be focused on the organisation's information requirements in support of its goals, critical success factors and plans.

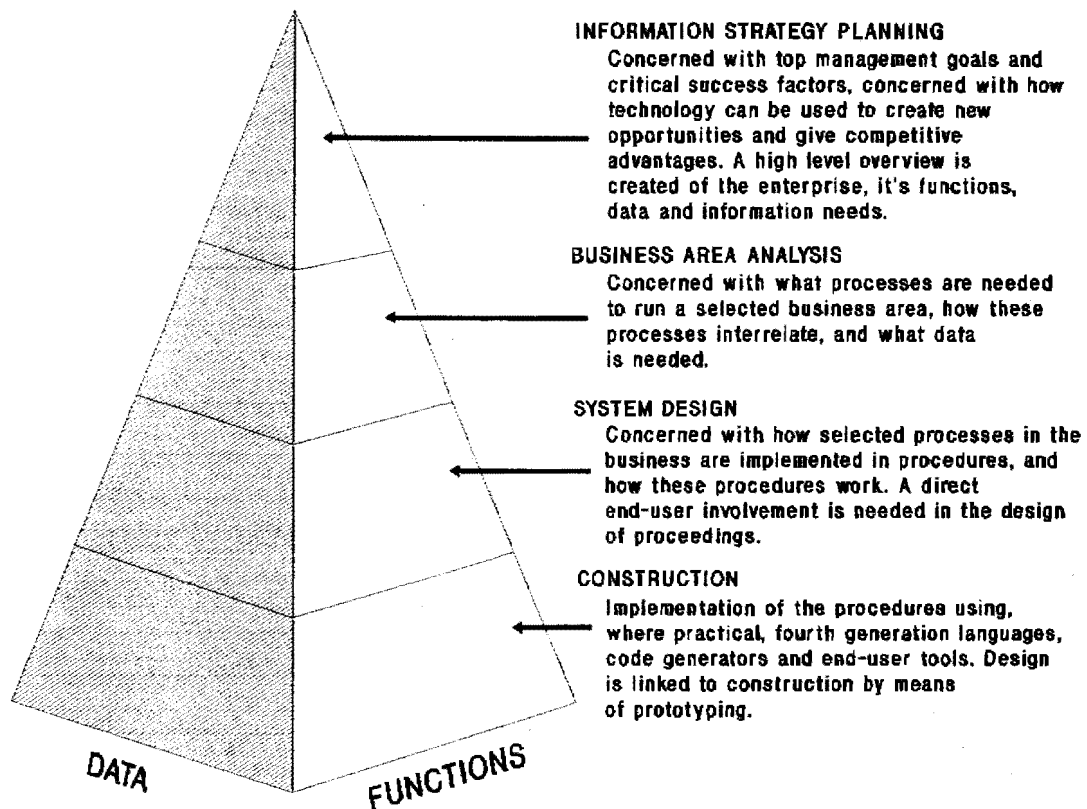
Porter and Perry (1984:40) describe the particular emphasis as organisation-wide and "top-down" in its approach, and as extending across all organisational components. They state that the top-down emphasis should take into consideration the organisation's structure, plan, model and long-range strategies, the organisation's association with the information that supports it and the individual components of each information system.

This particular "top-down" approach in determining the information requirements of an organisation and the provision of information systems using a corporate business approach within the ambit of IRM can be considered at four levels. This is shown by a "pyramid" in Figure 2.6 (Martin 1986:6).

These levels are:

- information strategy planning
- business area analysis
- system design
- construction

**FIGURE 2.6 AN INFORMATION ENGINEERING APPROACH TO SYSTEMS BUILDING**



Source: Martin, J. 1986. Information Engineering. Lancashire:Savant Research Studies. (page 10).

Martin (1986:6) observes that this approach, which is also known as an information engineering approach, requires a technique that allows



interested parties to gain a clear understanding of the business. This is best achieved through developing a model of the business.

The modelling process starts at the strategic or macro-level during the information strategy planning phase, which is then extended and normally refined during a particular business area analysis phase.

Martin (1986:11) explains that the phase that directs the information system focus is concerned with top management goals and critical success factors. It is also concerned with how technology can be used to create new opportunities or a competitive edge. During this phase a high-level overview of the enterprise, its functions, data, and information needs is created.

The business area analysis phase is concerned with what processes are needed to run a selected business area, how these processes are interrelated and what data is required (Martin 1986:74). The model that is compiled during the first two phases will be mapped into a particular systems during the design and construction phases of the systems development process.

Martin (1986:77) emphasises that information strategy planning is carried out to:

- investigate how better use of technology can enable an organisation to gain a competitive edge
- establish goals and critical success factors for the enterprise
- use critical success factor analysis for steering the organisation in the direction of goal attainment
- determine what information enables management to perform its work better

- prioritise the building of information systems in terms of their overall effect on the bottom line
- create an overview model of the organisation, its processes and information
- subdivide the overview model into business areas ready for business area analysis (level 2 of the pyramid)
- determine which business areas to analyse first
- enable top management to view its enterprise in terms of goals, functions, information, critical success factors and organisation structure.

The corporate, top-down approach, as depicted in Figure 2.6, illustrates the four phases of Information Engineering.

#### 2.4.2.3 Information technology architecture

It is not so easy to define the term information technology architecture (ITA) and there is no universal definition. It has been described by Earl (1989:97) as the technology framework that guides the organisation in satisfying business and the information needs of management. The Oxford Dictionary (1987:34) defines architecture as the art or science of designing and constructing buildings. Architect is defined, inter alia, as the designer of buildings and large structures.

ITA can therefore be defined as the compiled "blueprint", which is a design drawing that represents the business in terms of its information and technology, data, functionality and structure. Like the "blueprint" for the construction of a building the ITA (blueprint) serves as a framework for the construction of information systems.

Earl (1989:97) gives the following three main reasons why an ITA is necessary:

- an architecture provides a framework for resolving and reviewing technology choices over time, which is essential in the rapidly changing environment of today
- an architecture gives a structure for implementing the information technology needs of the organisation
- a technological model of the organisation is required to assist in the process of bringing IT strategy closer to business strategy

As systems become more complex and start spanning the organisation, it becomes more important to have a blueprint for the proper management of the information resources. One can just imagine how difficult it would be to construct a 20-storey building without a proper set of plans, whereas one could argue that a simple hut can still be constructed if adequate verbal specifications are given.

#### 2.4.2.4 Centralised information management

The concept "of user departments as it applies to data processing" has led to a gradual shift in control, away from senior management to individual users and departments, Ross (1982:1) explains.

With information systems and data spanning across the organisational lines (as a result of the implementation of data base systems) a high degree of coordination and integration on a corporate level is required, Ravindra (1987:1) emphasises. The developments in the area of data base systems and information management as a management discipline emphasise that a position similar to that of the other resource managers, like the financial controller, need to be created for the information resource portfolio (Ross 1982:1; Mach 1990:46).

Mach (1990:46) explains that to get information management under control, organisations are now appointing chief information officers. He points out that such appointments are required to align information management plans with business goals.

The creation of a separate and independent "information management" function not only facilitates the return of control over information to senior management, but also the recognition of information as a corporate resource of the whole organisation (Ross 1982:1; Curtice 1986:99).

Some important factors in successful information resource management have been highlighted. Other issues not mentioned include the risk involved in managing information incorrectly, the political realities of life – the people, power and policies of an organisation, user issues such as dissatisfaction and the advancing technology (Synnott & Gruber 1981:28).

## **2.5 SUMMARY**

Data processing has evolved over the years and the information era in which organisations are reaching towards maturity has now been entered. Today, systems are being developed to assist management in making decisions that have significant importance for the ability of the organisation to run profitably, compete, or even survive.

As organisations become more competitive to cope with the growing pressures of modern business, the providers of information have an increasingly important role to play in supporting the organisation in the achievement of their goals. Organisations can no longer run without sufficient information; managers can not manage without it. Yet, organisations are only just beginning to recognise the value of information resource management.

Now, already well into the 1990's, only those organisations that manage information as an important corporate resource will excel.

Against this background in satisfying the first objective of the study an overview of an IRM approach to information management and systems development, which included a discussion on the developments within EDP that gave rise to this IRM approach, was given.

## **CHAPTER 3**

### **SYSTEMS DEVELOPMENT WITHIN A CORPORATE DATA BASE ENVIRONMENT WITH SPECIFIC REFERENCE TO DATA RESOURCE MANAGEMENT**

#### **3.1 INTRODUCTION**

Data base technology is becoming increasingly important to data processing. McFadden and Hoffer (1991:29) assert that data bases have become a generally accepted technique for structuring and managing data in most organisations today.

Plagman (1982:1) states that the seventies witnessed the maturation of data base technology. Data base management systems (DBMSs), which were prone to software errors in earlier years, became comparatively stable and DBMSs and other data base technologies began to be employed as integral and important parts of a data base environment. Data base systems supported by data base management systems require a major reorientation to systems development and computer components since they tend to violate most of the basic premises on which businesses have traditionally constructed computer application and management systems (Techavichit 1979:1; McFadden & Hoffer 1991:30).

Techavichit (1979:14) records that one of the most important objectives of a data base system is to bring data resources and data users close together, thus making the data available when, where and how they are needed. This, in turn, explains how these concepts are intertwined and are in support of the overall objective of information resource management (IRM), in terms of getting the right information to the right people at the right time (this was discussed in the previous chapter).

In this chapter the concepts of a corporate data base environment, with specific reference to data resource management, will be introduced and

explored in terms of their impact on the systems development process. The purpose is not to dwell on the technical aspects but to focus on the areas that affect the management of the systems development process. The discussions will address those aspects of a corporate data base emphasis that support an overall IRM approach.

- section 3.2 will deal with the concept of data resource management covering the particular approach and the environment
- section 3.3 will address the issue of systems development within this new environment. In particular the impact of data base development on the systems development life cycle will be discussed
- section 3.4 will address systems responsibilities within the context of a formal data management approach
- section 3.5 will give a summary of the chapter

### **3.2 DATA RESOURCE MANAGEMENT**

Chapter 2 dealt with the information resource management concept. It was emphasised that IRM includes the management of information's underlying resources, such as its data raw material. Data resource management (DRM), for the purpose of this study, can therefore be viewed as a subset of IRM. This interpretation of DRM is supported by Sibley (1983a:2) who states that a more meaningful name for an IRM approach that encompasses DRM would be information and data resource management (IDRM), describing both types of entities and management emphasis. For the purpose of this study, DRM will be used to describe the particular way data has to be managed within the IRM perspective that was discussed in Chapter 2.

In this regard it is implied that data in itself should be treated as a critical corporate resource that needs to be managed pro-actively and to be strategically planned in terms of a properly defined data architecture that

forms part of a bigger information technology architecture. It is further implied that DRM will be managed from a co-ordinated corporate point of view and an information-engineering like approach will be followed to implement the data plan.

The following aspects will be addressed in this section:

- the data resource management philosophies : Section 3.2.1
- a data resource management definition : Section 3.2.2
- an overview of the organisational function of data management describing the differences between data administration and data base administration : Sections 3.2.3 and 3.2.3.1 respectively
- the data base environment in terms of its underlying concepts and components that supports a data management approach : Section 3.2.4

### **3.2.1 Data resource management philosophies**

To be successful in managing the data resources and in particular the data base technology of the organisation requires planning and engineering of the data pro-actively and controlling it to support the business of the organisation. The corporate data resources should be driven by the mission and goals of the business it supports (Martin 1986:32).

The data management approach adopted should reflect the particular corporate "style" or philosophy of the business being supported. It should also take account of the organisation's technical maturity. If data are perceived as a strategic asset, its management must take account of future information requirements (Synnott & Gruber 1981:281).

Once the exact business strategies and philosophies of the organisation have been established and understood, the resources required to manage and



control the data base should be identified and planned in detail and managed accordingly.

Data need to be managed on an organisation-wide basis and it is thus necessary to formally define the DRM function, which should be based on the following philosophies:

- recognition at a high level within the organisation that data is an important corporate asset and that it should be managed accordingly (Martin 1980:22)
- data are to be identified, defined and managed independently but in conjunction with the processes (applications) which use the data (Martin 1980:25). This is done in order to safeguard, protect and to ensure the integrity of the data and is achieved by analysing data and functions separately (Martin 1986:6)
- data are to be managed by a specialist group that will provide support for development of systems (Martin 1980:25)
- all the data components identified are to be defined in a data dictionary (Gilhooley 1985:5)

### **3.2.2 Data resource management defined**

Gilhooley (1985:4) defines DRM as the method of effectively managing the data resource of an organisation to provide the required information.

Porter and Perry (1984:40) describe DRM as an organisation-wide "management program" that is corporate in its aim and scope and strives to improve the integrity of the data.

In terms of the above definitions and for the purpose of this study, DRM will be defined as the corporate discipline of effectively managing the

organisation's data resources in support of the process of providing information with integrity to the organisation.

Sibley (1983:1) explains that the aims of DRM (in supporting the strategic direction of an organisation) are:

- making data accessible
- controlling data
- maintaining data integrity
- reducing development costs and reducing response time

Data resource management will be abbreviated as data management.

### **3.2.3 Data management function**

Many organisations now have at least 15 years of experience of data management and have well-established techniques for doing it as effectively as possible. Some have done an excellent job of data modelling and data administration. There are also those that have disastrously failed to achieve any overall coordination of data. The failure can be extremely expensive in the long run in terms of inflated data processing costs, failure to implement needed procedures and in lost business (Martin 1986:47; Martin 1989:73).

There are also many organisations that have not made a serious attempt at data management and allow systems to be built with ad hoc file design.

The trend within the context of DRM, is to separate the logical (business) view of data from the physical (data processing) view. In this regard a "data administration" function is established that is separate from the "data base" administration function, which is more concerned with the technical data processing issues (Gilhooley 1985:5).

### 3.2.3.1 Data administration versus data base administration

It is important to distinguish clearly between a data administration and a data base administration. Data administration is in general concerned with the "ownership" of data, its meaning, its relationships with other data and its global integrity. The data administrator is responsible for the high-level planning, modelling and coordinating of the organisation's data. It is a task of understanding the data required to run the business, and making diverse individuals agree about the definition and representation of the data.

It can further be viewed as the corporate service that assists in the provision of information by controlling and coordinating the definitions (meta-data format and characteristics) and the usage of reliable and relevant data. The scope of the function must be as wide as is required by the organisation in order to achieve the aims of cost-effective use of the organisation's data (Holloway 1988:7). Ravindra (1987:1) emphasises that the data administration function is now widely considered the key to more effective long-term data resource management.

Data base administration (DBA) is concerned both with the way in which the abstract data models are physically implemented in a computer system and with controlling the physical data base, that is the "custodianship" role over data. The data base administrator is concerned with a specific data base - the physical structure that has to be as effective and machine-efficient as possible (Martin 1986:47).

DBA is therefore primarily concerned with the technical implementation of the data base, the day-to-day operations of the data base and the policies governing its everyday usage. This entails the planning, coordination and administration of data bases as well as the DBMSs (Holloway 1988:19; Ravindra 1987:6; Sibley 1983:4).

There are a number of data management areas to which specific data administrator (DA) and data base administrator responsibilities should be

assigned. These responsibilities will be discussed in detail under data administration and data base administration, Sections 3.4.1 and 3.4.2, respectively.

#### **3.2.4 The data base environment**

There is a considerable amount of technical literature that addresses the data base environment and the operations of the DBMSs. In terms of this study only an overview of the significant concepts, as a framework for determining their impact on the systems development process, will be presented.

##### **3.2.4.1 Overview of the data base approach**

The move to a data base environment is not just a change in software; it is a total change in management.

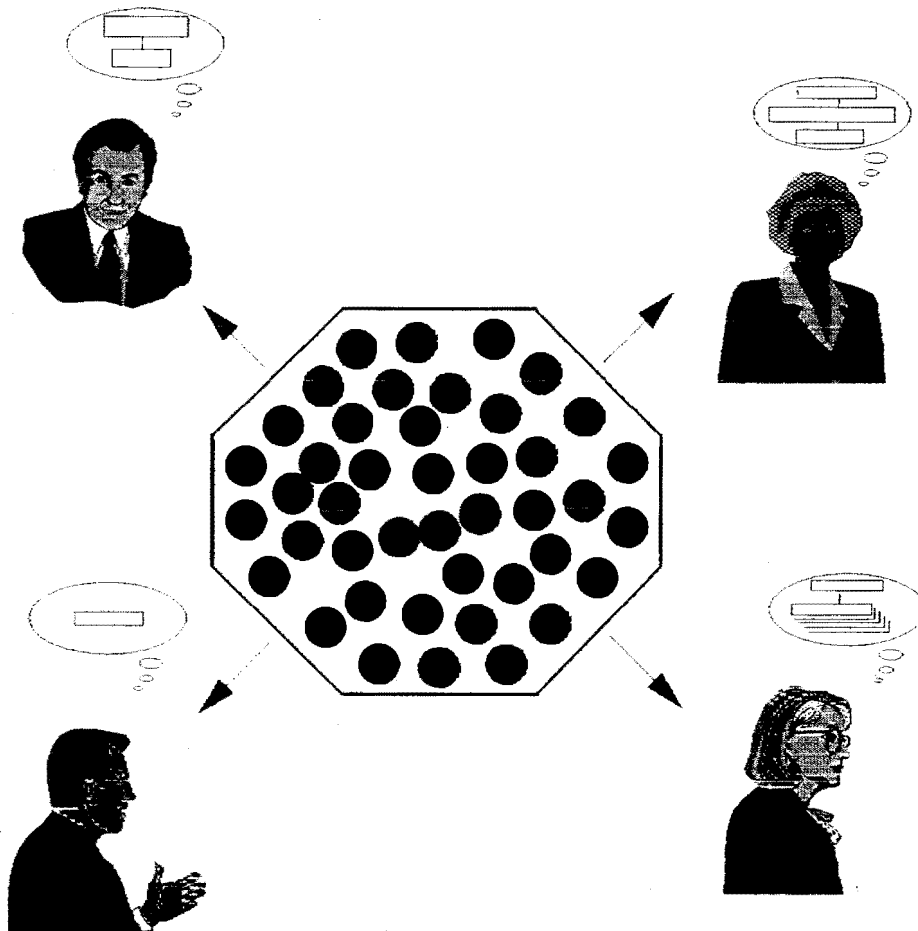
The purpose of data base technology is to provide the organisation with a set of tools that work together to manage the organisation's data resources (Meador 1988:1). The data base environment is based on the principles of data sharing and flexibility in data access. The goal of the data base is to allow users access to the data resources while maintaining control over security, integrity and auditability of the resources, Moss and Due (1988:1) record.

Data can exist in many different ways in a computer environment. In the traditional file environment a file or group of files are dedicated to an application. The data is normally physically stored in the form that makes it accessible for the application. It may be required to physically reorganise or sequence the file if another application program wants to use the data (AICPA 1983:1).

In a data base environment the data is organised in a more complex way. A software DBMS is used to structure the data. The DBMS allows multiple users to access the particular items of data, assembled in a central reservoir

in the way each program wishes to view the data, as is depicted in Figures 3.1 and 3.2 (AICPA 1983:1).

**FIGURE 3.1 THE COLLECTION OF DATA IN A DATA BASE**



Source: Martin, J., 1981. An end-user's guide to data base. New Jersey:Prentice-Hall. (page 51).

The data in a data base environment should be stored in such a way that they are independent of the applications programs in order to allow sharing. A common and controlled approach should be used in adding new data, modifying and retrieving existing data. The data should be stored in such

a way that it controls redundancy (Martin 1980:5). The concepts of data sharing and data independence are the main characteristics that distinguish the data base environment from the traditional application system environment.

AICPA (1983:4) describes the concept of data independence as follows:

*Data independence is accomplished when a single physical representation of the data is used to satisfy requests for data from multiple application programs. This independence is achieved by separating the application program from the physical data storage while maintaining the logical relationship through the DBMS. The DBMS maintains the definition of the logical views (sub-schemas) required by each application and constructs the sub-schema requested from the physical representation of the data.*

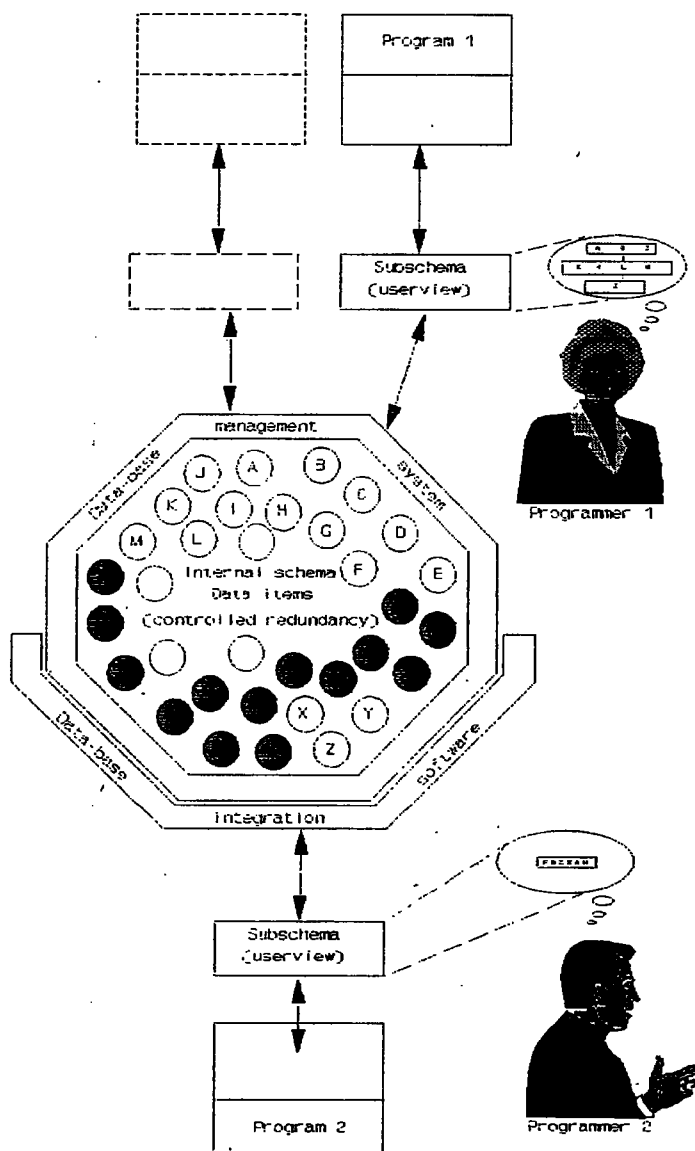
In terms of the above description of data independence, data sharing simply means that more than one application program require the same data elements for different purposes (AICPA 1983:4). The concepts of data independence and sharing are illustrated in Figure 3.2.

Martin (1976:78) explains that different application programmers (programs) have different views of the data. This figure depicts the sub-schema (schema) at the bottom with different user views (sub-schema's) at the top of the figure derived from the schema. The single physical representation of the data, refers to the overall schema. The term schema means the overall chart of all the data-item types and record types stored in the data base (Martin 1976:76). Information systems in a data base environment share data by means of a common base, which increases the accuracy and consistency of representation. In this sense a data base can be viewed as a shared collection of interrelated data designed to meet the needs of multiple types of end users Martin (1980:5).

Martin (1980:3) emphasises that a data base is not only shared by many users but is also perceived differently, as is depicted in Figure 3.2. In terms of controlling redundancy Duffy and Assad (1980:24) claim that it is not

suggested that only one "physical" data base is created in the organisation; nor that there should be no redundancy whatsoever. The emphasis is rather controlled redundancy.

**FIGURE 3.2 VARIOUS VIEWS OF DATA IN CONTEXT OF A DATA BASE ENVIRONMENT**



- Sources:
1. Martin, J., 1981. An end-user's guide to data base. New Jersey:Englewood Cliffs, Prentice-Hall. (pages 12&15).
  2. Martin, J., 1976. Principles of data base management. New Jersey:Englewood Cliffs, Prentice-Hall. (page 168).

#### 3.2.4.2 The components of a data base environment

There are ten basic components in a data base environment, as illustrated in Figure 3.3, namely:

- application programs
- computer-aided software engineering (CASE) tools
- data administration
- data base
- data base administration
- data base management system
- data repository / dictionary / directory system
- end users
- system developers
- user interfaces

#### APPLICATION PROGRAMS

McFadden and Hoffer (1991:26) explain that these computer programs are used to create and maintain the data base and provide information to users.

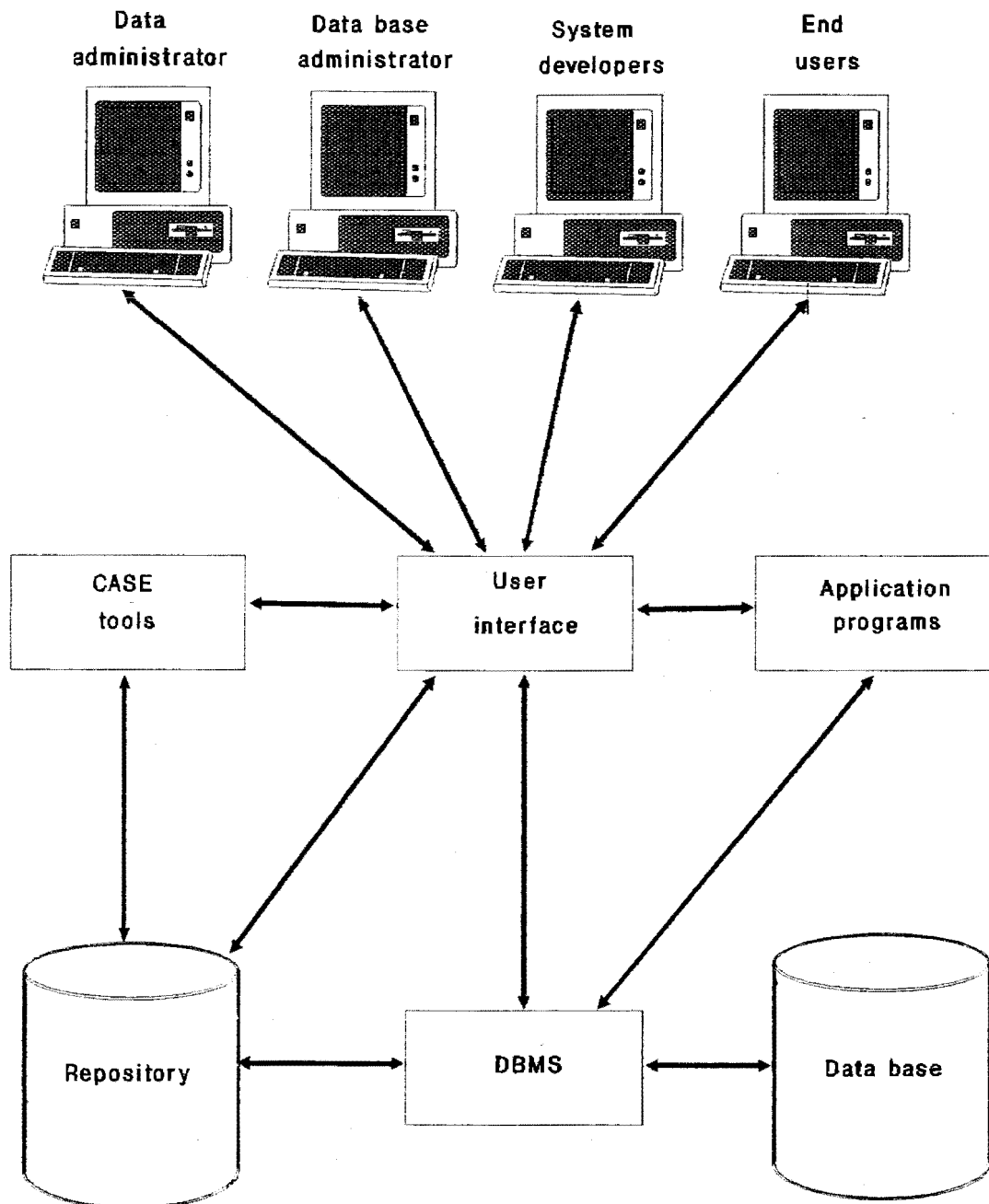
#### COMPUTER-AIDED SOFTWARE ENGINEERING (CASE) TOOLS

CASE tools are designed to support or automate the systems development life cycle (SDLC) phases. CASE tools are normally categorised into upper-CASE



and lower-CASE technology for the design and development of data bases and application programs respectively. (McFadden and Hoffer 1991:356)

**FIGURE 3.3 THE COMPONENTS OF A DATA BASE ENVIRONMENT**



Source: McFadden, F.R. and Hoffer, J.A. 1991. Database management. California: The Benjamin/Cummings Publishing Company, Inc. (page 27).

## DATA BASE

The data base itself can be viewed as a collection of data that is logically organised to support the requirements of a wide spectrum of users (Plagman 1982:2).

This description by Plagman is in line with Martin's definition adopted for this study, as described under Section 1.2.3.1 (Chapter 1) of this study.

There are a number of ways to structure or model this collection of data in a data base environment. The term data modelling is used to describe a way of organising data for a community of users and the different ways that the users like to have the data presented, Sibley (1983:7) explains. The more common types of data models are hierarchical models, network (CODASYL) models, relational models and entity – relationship models.

This study will only concentrate on entity–relationship modelling techniques which will be discussed in detail under Section 3.3.2.2 of this chapter and Appendix 1.

## DATA BASE MANAGEMENT SYSTEM

The DBMS is at the centre of a data base system. The DBMS consists of commercial software (and occasionally hardware and firmware as well) that provides access to the data base and the repository as depicted in Figure 3.3.

The DBMS manages the environment as opposed to reading and writing data to the data base. (Perry 1987:4).

## DATA REPOSITORY / DICTIONARY / DIRECTORY

All DBMSs have internal data directories that keep track of the data elements within the particular data base (AICPA 1983:12).

data definitions. It actually encompasses a large range of products. The tool may be used during one or more stages of the systems development life cycle for documenting and capturing information about structure, meaning and usage of data – data about data (meta-data). Holloway (1988:49) describes the differences between the 'activity' of dictionary systems as follows:

*Some are merely used for recording and subsequent enquiry, others are used as an integral part of the analysis, design, development and live running of systems. Some are based entirely around a particular database management system and some are actually integrated into the application database itself. The latter are referred to as integrated data dictionaries.*

The data dictionary/directory system in its broadest sense is known as a repository or knowledge base containing information about data definitions, screen and report formats, definitions of other organisation and system components that support the whole development life cycle, together with information about some operations activities such as change management, operating-systems management and network management. A data repository is thus a large and complex knowledge data base that collects data from several data-management and other software systems. (Butler Cox Foundation 1988:13)

For the purposes of this discussion the data repository can be viewed as a centralised knowledge base of meta-data (information about data).

## USER SYSTEM INTERFACES

The user interfaces include computer languages, menus and other facilities by which users within the data base environment interact with the various data base system components such as CASE tools, application programs, the DBMS and the repository (McFadden and Hoffer 1991:26).

## DATA BASE ADMINISTRATION, DATA ADMINISTRATION AND SYSTEM DEVELOPERS

The data base administration, data administration and the system development components of the data base environment will be discussed within the context of systems development in Section 3.4.

### 3.3 SYSTEMS DEVELOPMENT IN A DATA BASE ENVIRONMENT

#### 3.3.1 The role of systems development in general

The management objectives of systems development are constant, regardless of the technologies used in the implementation of information systems. The role of the systems development function is to utilise and effectively manage resources to achieve management goals (Plagman 1982:3).

The systems development function is responsible for software development projects that entail designing, implementing and maintaining information systems in support of the organisation's requirements (Plagman 1982:2). The systems development life cycle (SDLC) aids in managing and controlling development activities in that it breaks down the projects into manageable phases and tasks. It furthermore provides appropriate project control and management review functions (Gilhooley 1984:2). Projects are hierarchically structured into units of work called phases, activities, tasks and sometimes even smaller units of work.

The traditional SDLC consists of the following five phases, according to Plagman (1982:2):

- requirements analysis and feasibility
- systems design
- program design and coding
- testing and acceptance
- operations and maintenance

An information engineering approach to systems development, which was discussed in Chapter 2, divides the traditional SDLC into the following phases:

- strategic planning
- business area analysis
- design
- construction

The implementation and maintenance phases to this approach are implied. According to Martin (1986:12) the systems built under the information engineering discipline continue to evolve with business needs on an ongoing basis.

It is important to note that the four-stage information engineering process requires much more time to be spent on planning and design than on execution. In the traditional approach (as described above), the main efforts are towards the coding phase.

### **3.3.2 The impact of a data resource management emphasis on systems development**

The influence of DRM on the traditional role of the systems development process can be quite profound. Gilhooley (1990:4) states that the traditional emphasis on developing systems and, particularly operational systems, focuses on processing requirements rather than information.

In this sense the process-oriented systems design generally does not fulfil the strategic or even the tactical information requirements of the organisation. In most cases the information requests are not adequately addressed, because either the source data does not exist in the system or custom-built applications to handle the ad-hoc enquiries are too costly and time consuming.

Gilhooley (1990:4) adds that within a DRM approach to systems building the

"data-driven" emphasis overcomes this limitation of traditional systems development by focusing on data and information requirements during systems planning and building. Data is accommodated by means of its incorporation into various data models, which in turn, will be reflected in the application systems and enquiry routines written to convert this data into information.

A DRM approach to systems development results in:

- a data-driven emphasis to systems development
- a (separate) data base development life cycle

#### 3.3.2.1 A data-driven emphasis to systems development

The data-driven approach addresses problems inherent in the traditional systems development approach in two ways (Curtice 1990:3), namely:

- by making data the foundation of systems design and development in the belief that the definition and structure of data is more stable than specifications of reports and screen
- by specifying the data definition and structure independent of any individual application (or organisational unit), thereby facilitating the sharing of common data

Curtice (1990:3-4) adds that in terms of this data approach, the data base becomes an important, enduring product and should not be viewed merely as a mechanism to support an application. Data bases must be shared across applications and be made available to users as a resource for ad-hoc reporting.

It follows that the design of data bases within this environment cannot be based on detailed application specifications. What is required is a conceptual, or an enterprise-wide, data representation (model) that has to

be defined in advance of any particular development project. This model must also serve as an important input to the systems development planning exercise (Curtice 1990:6; Meador 1989:2).

Most observers view this particular approach as the planning phase of data management. The model that is produced during the planning phase is more commonly known as a data architecture (Butler Cox Foundation 1988:5; Appleton 1986:86-90; Meador 1989:2; Von Halle 1992:13; Gilhooley 1990:5).

This approach necessitates two separate but interdependent design processes that have to proceed in parallel – systems design and data base design. The system development process has to interact with the data base design process. This interaction can be defined in terms of a data base development life cycle, just as the systems design is defined in terms of a systems development life cycle (Plagman 1982:5; Martin 1989:4).

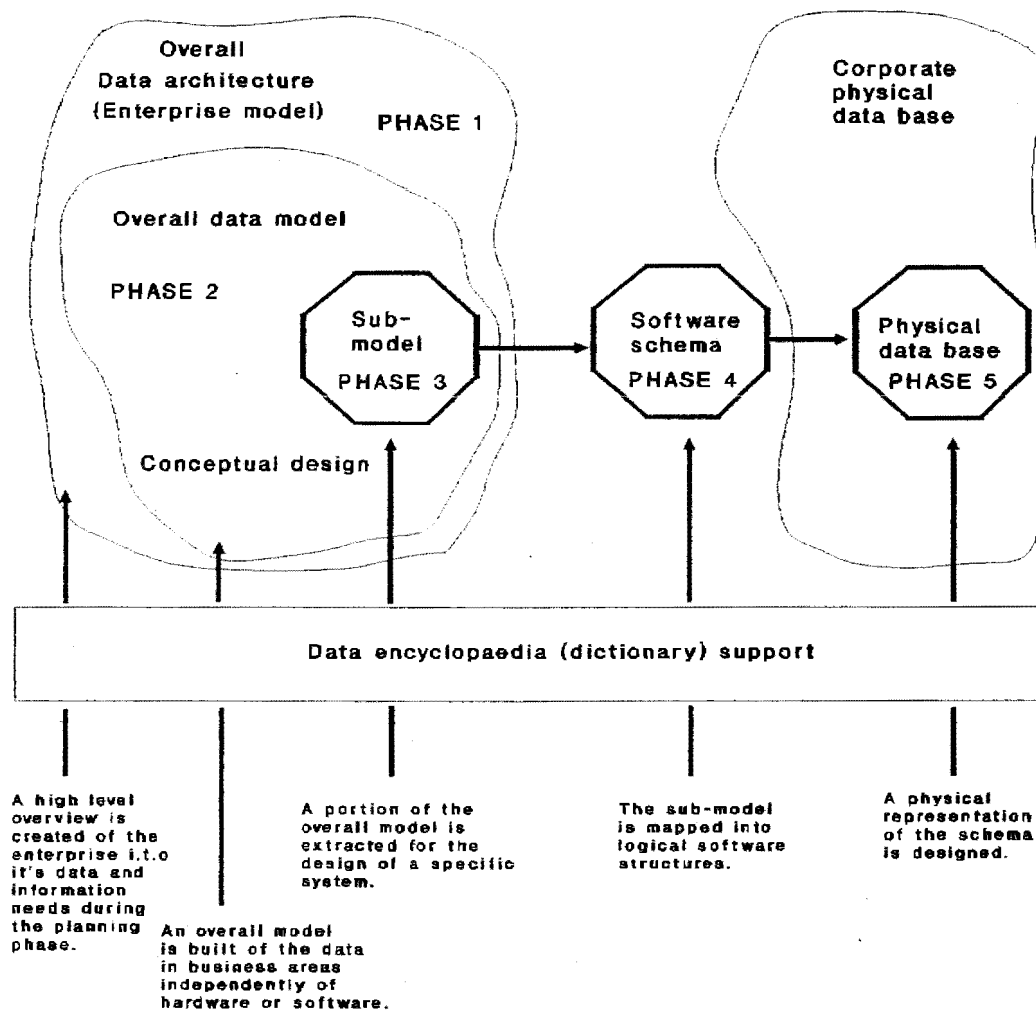
#### 3.3.2.2 A data base development life cycle

In its broadest sense, the data base development life cycle approach encompasses activities that range from the identification of end-user requirements to the final physical arrangement of data values.

Martin (1990:248, 252) divides the data base development life cycle approach (also known as a data base design methodology) into the following five major phases, as is depicted in Figure 3.4:

- overall data architecture compilation (data planning)
- overall data model compilation (conceptual design)
- sub-model identification and definition (detail conceptual design)
- logical data base design
- physical data base design

FIGURE 3.4 THE DATA BASE DEVELOPMENT LIFE CYCLE



**Sources:** Martin, J. 1990. Information engineering - Book II : Planning and analysis. New Jersey:Englewood Cliffs, Prentice-Hall. (page 253).  
 Martin, J. 1986. Information engineering. Lancashire:Savant Research Studies. (page 10).

#### OVERALL DATA ARCHITECTURE COMPILATION (DATA PLANNING)

Understanding the characteristics of the information that drives the business and identifying the source and location of that information (that is the raw data) are critical to an organisation's success. The data planning phase or data architecture compilation involves identifying the required data,



including the source and location and its characteristics and relationships with other data. The work performed in terms of data planning directly influences the other activities associated with managing the corporate data (Gilhooley 1990:5).

The data architecture can be viewed as the data "master plan" that identifies the key data items within an organisation and sets out the logical relationship between them. The data architecture describes the logical structure of the data used by the organisation that is essential to the development of integrated and flexible corporate data bases (Butler Cox Foundation 1988:2).

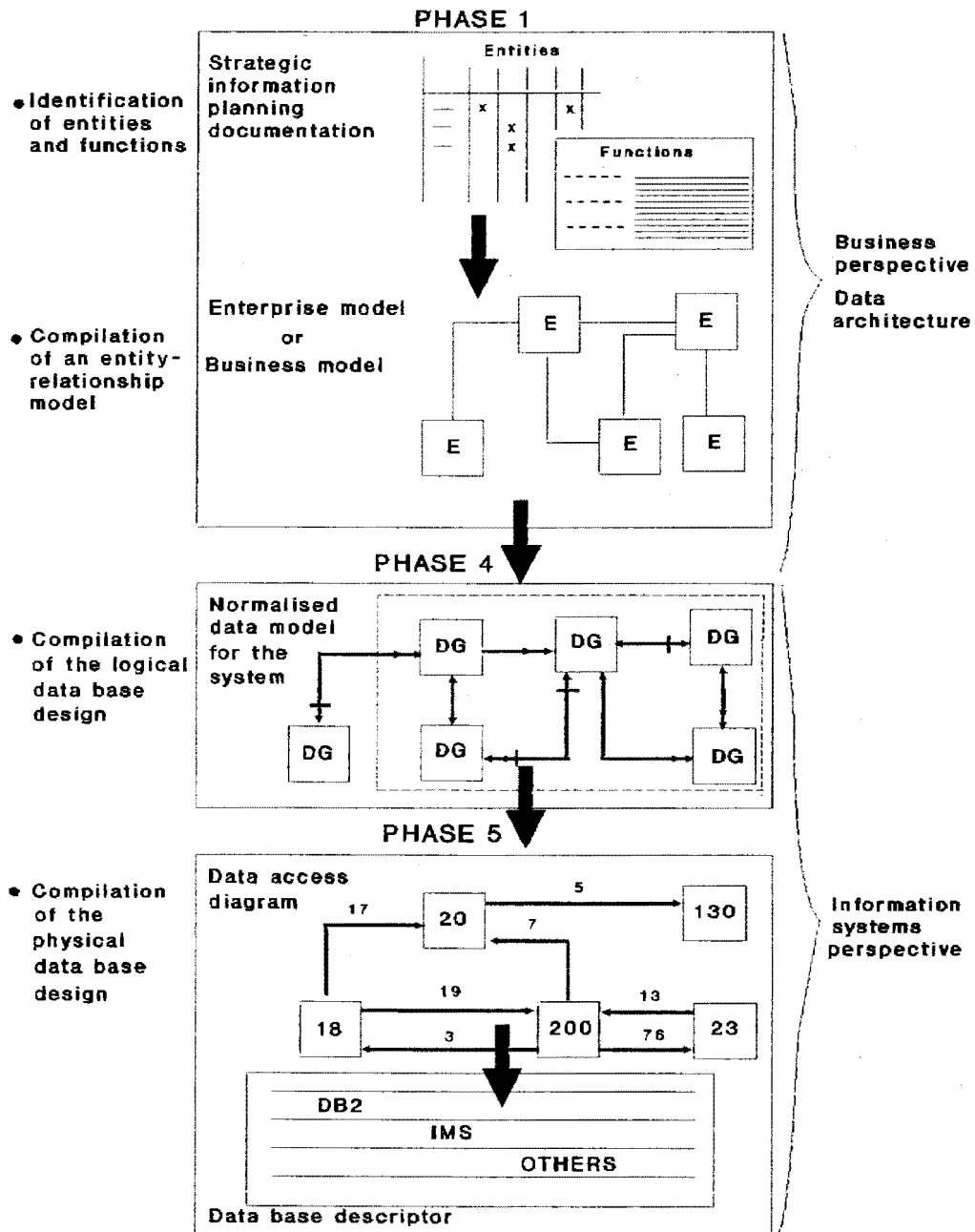
The process of town planning provides a useful analogy to the process of developing a data architecture as the Foundation (1988:2) explains:

*The purpose of town planning is to place different parts of the town (factories, houses, parks, shopping areas, and so on), in separate areas, but with easy access routes between them. The overall plan provides a framework for developing the town's areas and the relationships between them, and ensures that the detailed plans for each area can be integrated into a consistent whole.*

A data architecture performs a similar purpose by creating the framework for the development of corporate data bases and by providing the basis for the development and integration of individual applications in a data-base environment.

Appleton (1986:86) claims that an architecture must define enough detail to be meaningful – it must provide more of a means of seeing the forest than the trees. It must also provide a convenient, consistent way of resolving the forest into manageable subsets. In this regard the entity-relationship approach to compile the data model is still the best for establishing the planning structure. This approach implicitly produces two levels, one by defining entities – a high-level view for seeing forests, as it were – and one by defining relationships that exist between entities – a process that leads to focusing on the trees (Appleton 1986:86) as depicted in Figure 3.5.

FIGURE 3.5 A DATA ARCHITECTURE FRAMEWORK



NOTE: E = entities

DG = data (entity) groups

Phases 2 and 3 (of the business perspective) are not depicted

Source: Meador, J.G. 1989. A data architecture framework : The business perspective. In Auerbach data management. 21-01-01. New York:Auerbach (page 3).

Within the ambit of an information engineering approach the process of identifying entities will go hand-in-hand with the identification of functions during the strategic information planning phase, Martin (1986:6) further explains. The top block of Figure 3.5 illustrates this approach in terms of entities and functions listed during the strategic information planning phase. The second block represents the second level of data planning that now only concentrates on the entities. Business relationships between the entities are added during this phase.

The entity relationship model created in this way will now fulfil a two-fold role, namely it will firstly, act as a "blueprint" to direct the data management activities and, secondly, it will form part of the overall information systems architecture and systems plans of the organisation, the Butler Cox Foundation (1988:2) explains. Various information system models will now be derived from an entity-relationship model (second level model) within the context of an information systems and data base development environment, as depicted in Figure 3.5.

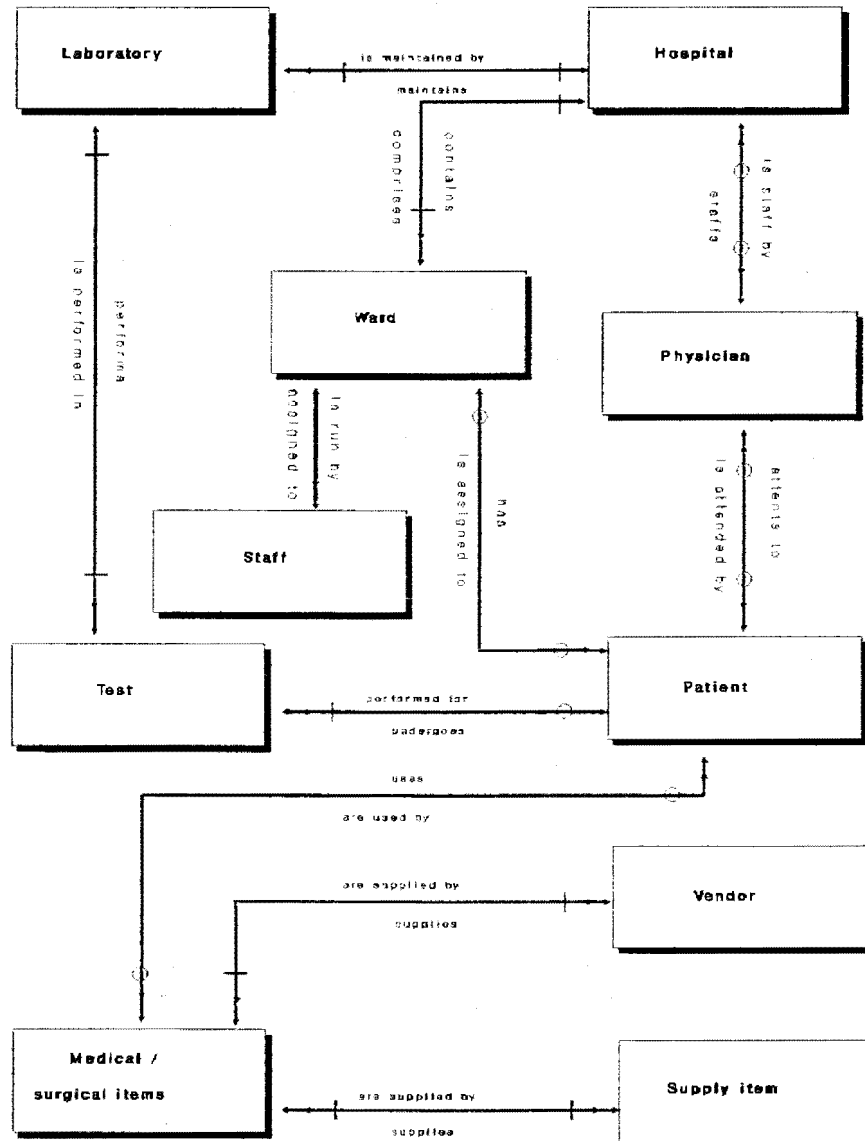
During the next phase data attributes will be added and normalisation will take place.

In explaining the approach an example of an entity-relationship model of a hospital is depicted in Figure 3.6. The figure depicts ten entities (hospital, laboratory, physician, etc.). These entities are linked with either "double" or "single" arrows explaining the "business" relationships that exist between the entities. A data modelling case study, in addition to the above mentioned entity-relationship model example is covered in Appendix 1.

#### OVERALL DATA MODEL COMPILATION (CONCEPTUAL DESIGN)

The overall corporate data model is derived from the data planning models (architecture) Ravindra (1987:4) states.

**FIGURE 3.6 AN ENTITY-RELATIONSHIP MODEL : MOUNTAIN VIEW  
COMMUNITY HOSPITAL**



Source: McFadden, F.R. and Hoffer, J.A. 1991. Database Management. California: The Benjamin/Cummings Publishing Company, Inc. (page 157).

Martin (1990:247-250) adds that the corporate data model in terms of a top-down (information engineering approach) will be derived from the entity-relationship model created during the information strategy planning phase, as is depicted in Figure 3.5. It follows that the data overview created during

the strategy planning phase is actually a pre-phase of the compilation of the data model as Martin (1990:250) explains:

*Top-down design creates an entity - relationship diagram which gives an overview of the data entities that the organisation should work with, but not fine detail of the data structures. It should look at the future evolution of the enterprise and its use of information. Detailed design then extends this overview and creates a normalized data model.*

Martin (1986:42) adds that although the detailed design phase expands on the entity relationship model the model is not completed yet. The emphasis is to make the model as stable as possible, but without going into too much detail. The "detailed" descriptions of the data items will be added during the sub-model definition phase.

A data repository or dictionary must be utilised to document and rationalise data descriptions and information about data models already during this phase, the Butler Cox Foundation (1988:3, 6) advises.

Plagman (1982:5) aptly calls this phase the "global conceptual data base design", which results in an overall corporate data structure diagram (model) and its supporting documentation.

#### SUB-MODEL IDENTIFICATION AND DEFINITION (DETAIL CONCEPTUAL DESIGN)

During this phase a detailed model for the specific systems development project, based on the overall data model, is developed. The particular model can be viewed as a sub-model of the overall data model (Ravindra 1987:4; Martin 1990:247). Martin (1990:247) claims that it is like an architect creating an overall plan of a house (overall data model) first and then drawing one room (sub-model) at a time in detail.

This phase results in a clear definition of the data sub-model in terms of the

specific system boundaries and its supporting documentation (Plagman 1982:5; Martin 1990:247). During this phase a detailed analysis of the data is carried out. The detailed descriptions of the data items (attributes) used in the data models are collected and synthesised into a fully normalised data model for the system (Martin 1986:44; Butler Cox Foundation 1988:6). The data model produced thus far is independent of the software or hardware that will be used, Martin (1990:252) adds. The data dictionary will also be updated during this phase to the required level of detail Meador (1989:8) adds.

Some observers believe that the sub-model identification and definition phase actually forms part of the logical data base design phase (Meador 1989:6; IIARF 1991b:5-21).

#### LOGICAL DATA BASE DESIGN

When data bases are implemented, the data models produced are translated into logical software structures such as relational tuples. These logical software structures are created in terms of the particular data base management system. This results in a logical data base design. (Plagman 1982:5; Martin 1990:252).

The purpose of the logical data base design is to map the data model into a schema that can be processed by a particular DBMS. DBMS-processable schemas are developed using the data description language for the DBMS to be used. Logical design is considered an intermediate step between "conceptual" and physical design McFadden and Hoffer (1991:170) explain.

A variety of decisions have to be made before the design can be physically implemented. These decisions depend on aspects such as how the data are likely to be used, the usage paths through the data base, their volumes of use and response times required (Martin 1986:46).

## PHYSICAL DATA BASE DESIGN

During this phase the physical representation of the logical data base structures are designed. Storage, allocation, buffer and block sizes and physical placement options are addressed and this results in a fully described data base design, coded and ready for data base loading procedures as McFadden and Hoffer (1991:170) explain:

*In this stage the logical database structures (normalized relations, trees, networks, and so on) are mapped to physical storage structures such as files and tables. Indexes are specified, as well as access methods, record blocking, and other physical factors. A major objective of physical design is to provide adequate performance for user applications in terms of response times, throughput rates, and so on. Also, physical database design is concerned with database security, integrity, backup, and recovery.*

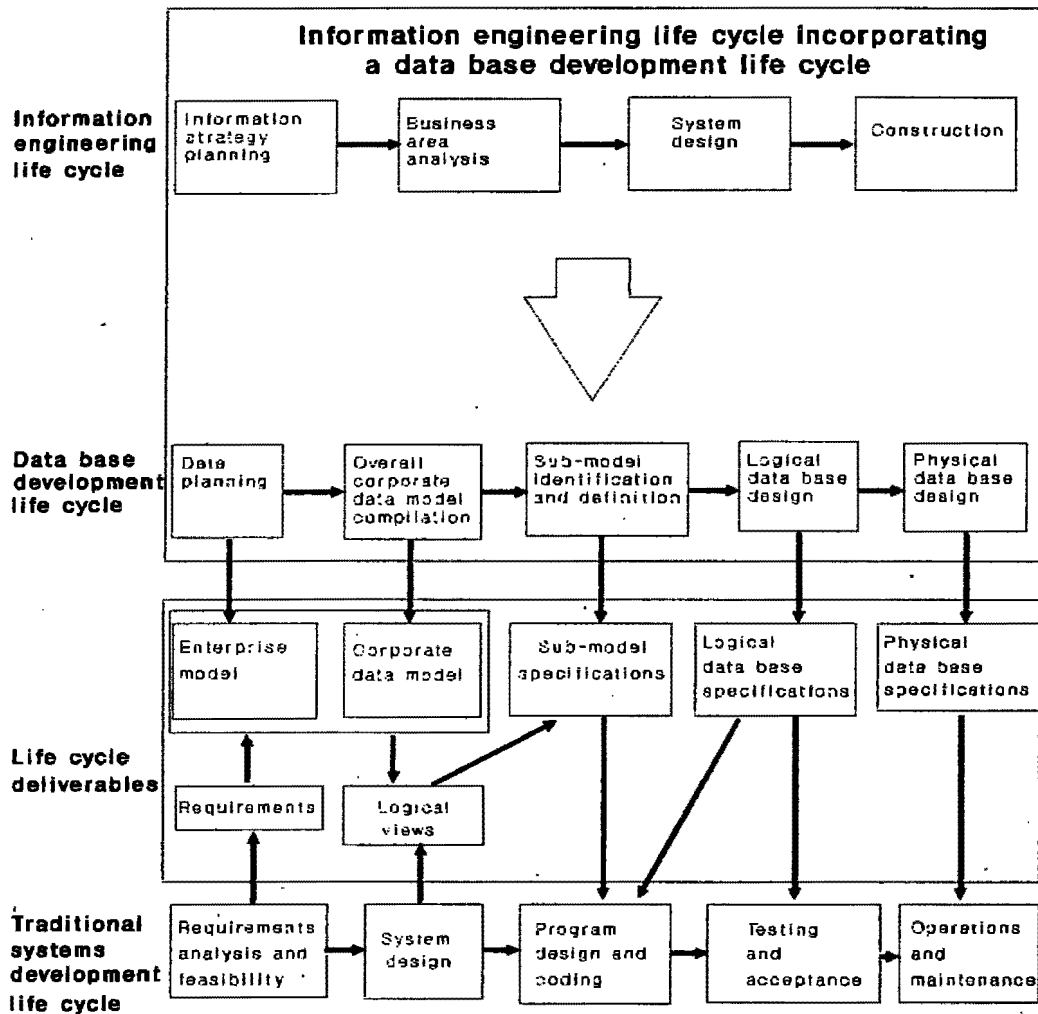
### **3.3.3 The interaction between the data base development life cycle and the traditional systems development life cycle phases**

It was explained in the previous section that the DRM approach in addressing an organisation's information needs has a dramatic impact on the systems development discipline. Figure 3.7 illustrates the manner in which the traditional SDLC is effected by the introduction of data base technology.

The interaction between the traditional systems development life cycle and the data base development life cycle is indicated and special reference is made to the information engineering approach to systems building. Figure 3.7 depicts, in the top box of the picture, the information engineering development phases that incorporate a data base development life cycle. The box in the middle indicates the data deliverables in terms of the data base development and traditional systems development life cycles respectively.

The discussions that follow will concentrate on the interaction of the SDLC phases pertaining to the data base development life cycle phases.

**FIGURE 3.7 INTERACTION BETWEEN THE DATA BASE DEVELOPMENT LIFE CYCLE AND SYSTEMS DEVELOPMENT LIFE CYCLE**



- Sources:**
1. Plagman, B.K. 1982. Systems development in a data base environment. In Auerbach data base management. 23-02-01. Auerbach. (page 7).
  2. Martin, J. 1986. Information engineering. Lancashire:Savant Research Studies. (page 10).

### 3.3.3.1 Requirements analysis and feasibility phase

This is the first step in the traditional systems development process. During



this phase efforts focus on user-oriented and conceptual aspects of the design. This phase is to establish the objectives of the systems, determine its scope, and review its technical, operational, and economic feasibility. The main objective is to obtain an understanding of the user's problem; an appropriate computerised solution; a report for management, complete with a cost-benefit analysis in sufficient detail for an intelligent judgement to be made as to the desirability and justification of the project; and the allocation of the necessary resources for its undertaking.

Within the context of a data base environment the traditional activities have to be coordinated with the data base development life cycle (Ravindra 1987:4). An application project deals with a subset of the overall (conceptual) data model and refines the applicable section of the data structure in accordance with specific user requirements by providing additional conceptual insight (Plagman 1982:6).

Gilhooley (1990:8) adds that within the ambit of this new environment and data management approach systems should be defined within the context of the information strategy, which in turn should emanate from the development of the corporate data model that precedes this phase.

An aspect of particular concern pertaining to the feasibility of an application system is whether it can utilise the data bases defined in the global data model and whether the particular data base technology (for example a DBMS) of the organisation must be employed. The requirements and feasibility report must formally address the data requirements. The report should also address the user requirements that are related to the access, storage, and maintenance of information in the data bases (Plagman 1982:7).

#### 3.3.3.2 Systems design phase

This phase is concerned with the detailed analysis that leads to the point where programming and procedure development can begin. The modules that make up the system are defined in precise detail in terms of detailed systems

flow, file descriptions, input/output procedures and forms requirements during this phase.

Nasuti (1987:7) states that the design is one of the most important phases of the systems development life cycle. The following design specifications are realised during this phase:

- the designation of all output required by the users
- the determination of the input needed to achieve the required output, including format and type
- the processing needed to achieve the objectives of the system
- the establishment of the type of file organisation

Within the ambit of the data base development approach the data sub-model has to be specified in conjunction with the functional requirements (Martin 1986:82). Plagman (1982:7) stresses that the logical views of each application must be carefully detailed by enumerating the transactions required to support the business functions in question. It can be viewed as an analysis that identifies the data elements and data relationships necessary to support the application. The retrieving, updating, and modifying requirements need to be established. The data sub-model definition should form part of the formal deliverable representing the logical views of the application as depicted in Figure 3.7.

#### 3.3.3.3 Program design and coding phase

Gilhooley (1990:8) explains that this phase entails the development of the computer programs and manual procedures of the total system. The previous phase (system design) resulted in the design of each element of the system, detailed and accepted by systems management. Programming now proceeds rapidly from this point, giving real visibility to the system for the first time.

Program flowcharts, coding, and unit testing proceed from detailed design specifications. The tasks in this phase normally require knowledge of the data structures and access methods used in the data base design.

Plagman (1982:8) claims that it is of the utmost importance that the system development staff employ procedures and techniques to promote data independence and the overall integrity of the data base during this phase. Special care should be given when writing programs that update, modify, and/or delete data from the data base. In addition, attention should be paid to edit and validation controls, DBMS return code processing and other error processing procedures.

The system's data requirements are expressed by means of the logical data base design as discussed earlier.

#### 3.3.3.4 Testing and acceptance phase

This phase entails the testing of the total system to ensure that it is acceptable to the users and meets all the operational requirements.

In terms of data being shared across specific application boundaries (within the context of a data base environment) it requires an additional level of testing – data base integration testing. The newly developed system needs to be tested against a test data base, along with other application programs that are in production mode (Plagman 1982:8; Gilhooly 1990:4).

#### 3.3.3.5 Operations and maintenance phase

The system is now implemented and in production to fulfil the initial requirements. This final phase is an ongoing phase in which the system will be maintained. (Mendus 1986:39).

Martin (1986:47) points out that maintenance is an ongoing concern and in terms of the data base approach it is important to keep the data models up to

date in respect of data changes.

Maintenance is critical in this new environment and must again be carried out in terms of the overall data resource management principles of data accessibility, control and continuity. It also implies that a new control structure is required for the organisation, namely new administration functions and in particular data administration (DA) (Sibley 1983:2,10).

### **3.4 SYSTEMS DEVELOPMENT RESPONSIBILITIES WITHIN THE CONTEXT OF A FORMAL DATA MANAGEMENT APPROACH**

Gilhooley (1990:6) claims that a chief information officer (CIO) should be in charge of the information systems department. The CIO should be responsible for formulating an information strategy that includes all systems development, computer operations and communications planning and operation.

Gilhooley (1990:7) lists the following four organisational groups that should be involved in the development of systems in a data base environment in addition to the users:

- data administration, which is responsible for developing and maintaining the integrity of the logical view of the corporate data base and, typically, the custodianship of the data dictionary
- data base administration, which is responsible for the physical representation of the data base (custody), the performance of the production data base and data security
- systems development, which is responsible for developing the logic and building the system that turns data into information
- computer operations, which are responsible for systems operations and monitoring

**TABLE 3.1 RESPONSIBILITIES AND INTERACTIONS BETWEEN ORGANISATIONAL GROUPS INVOLVED IN SYSTEMS DEVELOPMENT WITHIN THE CONTEXT OF A FORMAL DATA MANAGEMENT APPROACH**

PHASE	PHASE DESCRIPTION	INTERACTIONS
<b>Pre-phase: Overall system definition (Planning)</b>	Preliminary identification of the systems to be developed. System definitions are produced as a result of developing the information strategy.	System definitions emanate from the development of an information strategy and corporate data model. This strategy is developed by CIO in conjunction with the managers of systems development and data administration.
<b>Requirements analysis and feasibility</b>	Detailed study on the cost and benefits of developing the new system.	This development phase is carried out by the systems development group. Approval to proceed to the next phase must be obtained from the DP steering committee.
<b>System design</b>	Expansion of the general design to the point where programming and procedure development can begin.	Systems development manages this development phase. However, a data analyst is assigned from data administration for as much time as is necessary to define the data elements and structures and to modify and document the test version of the data dictionary and corporate data model. Approval of the detailed design must be received from data base administration, computer operations, and data administration before proceeding to the next phase.
<b>Program design and coding</b>	Development and testing of all computer programs and manual procedures (i.e. develop the total system).	This phase belongs almost exclusively to systems development. However, there will probably be periodic interaction with the data analyst if changes to data definitions are required.
<b>Testing and acceptance (including implementation)</b>	Independent testing of the entire system to ensure that it is acceptable to the users and meets all operational requirements.	Data base administration, computer operations, and the users are involved in this phase. All parties must sign off before the introduction of the system into production.
	Controlled migration of the system from the test environment to the production environment.	The move from test to production involves computer operations and data administration. Computer operations moves all object and source code into production. Data administration updates the production data dictionary and corporate data model.
<b>Operation and maintenance</b>	Ongoing maintenance of the system.	Data administration coordinates all aspects relating to data changes in conjunction with system development's responsibilities for system changes and maintenance

**Source:** Gilhooly, I.A. 1990. Managing the corporate information resource. In Auerbach data base management. 24-02-10. New York:Auerbach. (page 8).

Table 3.1 presents an overview of the responsibilities and interactions of these groups during the systems development life cycle and its subsequent operation in a production environment within the context of a formal data management approach. Gilhooley (1990:7) points out that it is very important that the various roles are understood and clarified.

In terms of this study further discussions will only concentrate on the respective data administration and data base administration responsibilities.

### **3.4.1 Data administration**

Many authoritative observers support Martin (1986:47) in his claim that the data administration function should be responsible for:

- determining the information needs of the organisation
- creating and maintaining the corporate data models, ensuring that they are as stable as possible
- obtaining agreement among users about the definitions and format of data items
- ensuring that systems builders conform to the data models as far as possible
- resolving conflicts about incompatible representations of data

#### **3.4.1.1 Determining the information needs of the organisation**

Martin (1986:3) claims that the information needs of an organisation can only be determined through an information engineering type approach to systems building, which was discussed in detail in Chapter 2 (Section 2.4.2.2) and Section 3.3 of this chapter. He points out that this "top-down" approach starts during the information strategy planning phase during which an

overview (architecture) of the organisation's data is created.

Gilhooley (1990:6) emphasises that a properly constructed architecture places the systems to be developed into a proper business perspective and also notes that the architecture is instrumental in the preparation of the information systems department's strategic plan.

Appleton (1986:86) and many other observers (Ravindra 1987:2; Holloway 1988:12-13 and Durell 1984:6) also agree that the data administrator (DA) should have global responsibility for the planning and policy development of the corporate data resource, which should include the development of a strategic information systems plan.

Ravindra (1987:2) adds that to ensure control over the evolution of information systems in an enterprise requires:

*strategic information systems planning, which is the process of determining objectives, resources, and policies that will govern information in an enterprise. As the function chartered to manage information as a corporate resource, DA must give strategic planning its close and ongoing attention.*

Holloway (1988:13) elaborates on this point by saying that the DA function should also be responsible for educating the organisation on the importance of data and creating an awareness of the value of data, and should also spread the knowledge of what data exists.

#### 3.4.1.2 Creating and maintaining the corporate data models and ensuring that they are as stable as possible

Gilhooley (1990:6) and Martin (1986:19) claim that the data administration function is responsible for building and maintaining the corporate data model.

Ravindra (1987:2) and Appleton (1986:90), in line with this view, emphasise

that the DA function should develop the data model (also called a global conceptual model) for the organisation. They further emphasise that the DA should also coordinate and integrate the user data models.

Holloway (1988:14) claims that the DA must assist in the process of business analysis for the production of data models.

#### 3.4.1.3 Obtaining agreement among users about definitions and format of data

Appleton (1983:226) states that one of the main problem areas of data management is the aspect of establishing consensus amongst all the concerned parties about the "meaning" of data. The DA function has to liaise with the end users and the data processing team to define each data element – its size, format, usage and other pertinent information. This information should be properly documented in a data dictionary (Ravindra 1987:5, Ross 1982:7).

In this regard Durell (1984:6,7) points out that the DA is concerned with the management of information about data (also known as meta data) stored in the dictionary.

#### 3.4.1.4 Ensuring that systems builders conform to the data models as far as possible

Martin (1986:42) explains that within the ambit of an information engineering approach to systems building the systems design phase proceeds by utilising the data model created during the business area analysis (BAA) phase that was discussed in Chapter 2. From the overall data model created sub models are extracted for the design of specific systems, as explained in Section 3.3.2.2.

To ensure that systems builders conform to the data models, requires that:



- high-level data models are in place (Butler Cox Foundation 1988:7; Appleton 1985:111; Appleton 1983:230; Martin 1986:42)
- a "policy" for a single source of data is promoted and enforced (Holloway 1988:14; Ravindra 1987:2; Sibley 1983:2)
- standards, procedures and controls necessary for data management are in place. (Ravindra 1987:5; Holloway 1988:12-17)
- the DA coordinates and controls the system builders' data models and data documentation in that he is made accountable for the data documentation (Ravindra 1987:2; Sibley 1983:2; Martin 1986:50)
- a data dictionary, under the control of the DA, is utilised and stores all the information about the data (meta-data). Data dictionary standards must be in place and should be enforced (Ravindra 1987:5; Holloway 1988:14; Sibley 1983:7)

#### 3.4.1.5 Resolving conflicts about incompatible representations of data

There is usually substantial argument about the representation of data Martin (1980:63) comments. The DA must establish agreement about common definitions and data structures before the data can be put into a shared data base to ensure that the multiple users are served as effectively as possible. The DA must have authority to arbitrate in disputes and establish consensus in the various aspects of data usage (Holloway 1988:15; Appleton 1983:225; Sibley 1983:3).

#### 3.4.2 Data base administration (DBA)

The data base administrator's main function include:

- the design of the data base

- maintaining data integrity, security, accuracy, and completeness
- coordinating computer operations
- monitoring and improving systems performance
- providing administrative support

#### 3.4.2.1 The design of the data base

Ravindra (1987:6) and Durell (1984:7) explain that the DBA is responsible for the actual technical data base design that provides the foundation for the integration of corporate data resources residing in the computer. Durell (1984:7) emphasises that one must remember that the data administrator is normally responsible for the logical data base design whereas the data base administrator is usually responsible for the technical data base design.

The data base administrator normally modifies the logical design where necessary to increase performance of the DBMS and the hardware in designing the technical data base. The DBA must consider the current and future processing, security, integrity, performance and data requirements for the data base design to ensure that the various systems' and user's requirements can be met on a timely basis (Holloway 1988:20; AICPA 1983:11).

#### 3.4.2.2 Maintaining data integrity, security, accuracy and completeness

The Report of the Joint Data Base Task Force, published by the American Institute of Certified Public Accountants (AICPA 1983:11), lists as part of this responsibility, the issues of who may access data and how the access is accomplished; protection of the data base against inaccurate, invalid, or missing data; securing the data base from unauthorised access and destruction and total recovery in event of loss.

Perry (1987:2) claims that although the DBA is responsible for the safekeeping of the data delivered to the data base, the DBA cannot assume responsibility for the accuracy, completeness, and authorisation of transactions that access the data base. He correctly points out that the responsibility for ensuring the integrity of a data base is a custodial one only. He adds that the responsibility is five-fold, namely to:

- safeguard the data delivered to the base
- protect the data base system from unauthorised access
- ensure that the data base can be reconstructed in the event of problems
- establish the required organisational structure that provides the necessary checks and balances
- provide the end users with controls that ensure the integrity of actions and reports compiled that utilised the data base

#### 3.4.2.3 Coordinating computer operations

This function includes assigning responsibility for the physical computer resources and monitoring their use to minimise access time, central processing unit (CPU) cycles and data base reorganisations and to maximise data security, disk space utilisation and the strengths of the data base management systems (AICPA 1983:11; Durell 1984:12).

#### 3.4.2.4 Monitoring and improving system performance

Performance becomes a common issue when several users share a corporate data base and all have to go through one data base management system to access the data. The DBA must set standards, procedures, and performance-monitoring mechanisms. The DBA's responsibilities include monitoring performance, recognising deficiencies and taking corrective steps

to improve performance (AICPA 1983:12; Ravindra 1987:6).

#### 3.4.2.5 Providing administrative support

The DBA's role furthermore requires coordination and liaison with DBMS vendors and assessing new releases of DBMS software and establishing technical standards and training for the data base environment (Holloway 1988:20; AICPA 1983:12).

Holloway (1988:20) concludes that the DBA function is a "clearing house" for the collection, classification and skills required for the success and maximum benefit of the data base system.

### 3.5 SUMMARY

In attempting to satisfy the first objective of this study, namely, to introduce the concepts of data resource management within a corporate data base environment and to determine its effect on the systems development process, an overview of the relevant data management and data base concepts were given.

It was explained that the notion of common data shared by numerous users is fundamental to the concept of a data base environment. It was further explained that the concept of data base technology is very much intertwined with the concepts of information and data resource management. It was stressed that, for the purpose of this study, a corporate data base approach to systems development implies formal data resource management. A data resource management emphasis, in turn, implies the management of the data resources within the context of a formal information resource management approach.

Data resource management (or data management) was explained as the organisation-wide function required to manage data as a critical resource and ensuring integrity in support of the process of providing meaningful

information. It was illustrated that a data management approach to systems development differs radically from the traditional (file oriented) systems development process in that two separate but interdependent design processes now have to proceed in parallel - systems design and data base design.

The effect of this new approach on systems development was highlighted as it pertains to the traditional systems development phases.

## **CHAPTER 4**

### **THE IMPACT OF A FORMAL DATA MANAGEMENT APPROACH WITHIN THE CONTEXT OF A CORPORATE DATA BASE ENVIRONMENT ON SYSTEMS DEVELOPMENT AUDITING AND AUDITING IN GENERAL**

#### **4.1 INTRODUCTION**

The previous two chapters dealt with the first objective of this study. It was explained that data resource management, as an organisation-wide discipline, is necessary to ensure that data with the required integrity is available and structured in accordance with the organisation's information objectives. It was also illustrated that data has to be managed independently but in conjunction with the systems development life cycle in terms of its own data base development life cycle.

It is against that background that this chapter will now address the second objective of this study, namely, to determine the impact of a data management approach (within the context of a corporate data base environment) on systems auditing and in particular the systems development auditing process.

The discussions of this chapter will concentrate on the data component of information systems and an audit approach that is of mutual concern to all types of auditors: compliance with appropriate internal controls. Substantive auditing tests of results, which is primarily a concern of external auditors, are not covered. Mair, Wood and Davis (1986:3) propagate that the emphasis in this particular approach is to predict the reliability of information systems based upon the internal controls been evaluated.

Within the context of the scope of this study, in addressing the data component of information systems, the discussions will concentrate on the aspect of data reliability which is more commonly known as data integrity.

The discussions will further focus on the logical integrity aspects of data that refers to the modelling of data as well as ensuring that models are physically correctly implemented and that the meaning of data in terms of "meta-data" is maintained.

The view is expressed that an information system will provide reliable information when appropriate controls which maintain the integrity of the data (logical and physical) as well as the processes are in place.

In this chapter it is proposed to:

- firstly, discuss the role of the auditor pertaining to the evaluation of the system of internal controls. Specific reference is made to the process of evaluating controls during systems development (Section 4.2)
- address the aspect of ensuring the reliability of information systems with specific reference to data integrity and data management (Section 4.3)
- discuss the effect of formal data management, in the context of a data base environment, on the organisational control structure and the audit process (Section 4.4)
- lastly, determine the impact of a formal data management approach on the traditional systems development audit process (Section 4.5)

## **4.2 THE ROLE OF THE AUDITOR IN THE EVALUATION OF THE SYSTEM OF INTERNAL CONTROLS DURING SYSTEMS DEVELOPMENT**

Porter and Burton (1971:5) define auditing as:

*...the examination of information by a third party other than the preparer or the user with the intent of establishing its reliability and the reporting of the results of this examination with the expectation of increasing the usefulness of the information to the user.*

In this sense the objective of auditing is to increase the reliability of information where reliability may be defined as congruence between the messages transmitted and the realities being described, claim Porter and Perry (1984:7).

Within the ambit of this broad objective, many types of information may be audited by different types of auditors. The external (attest) auditors, for example, are usually involved in rendering opinions on financial statements whereas internal (management) auditors are typically involved in independently appraising business activities as a basis for serving management (Porter & Perry 1984:8).

The Canadian Institute of Chartered Accountants (CICA 1975:159) explains:-

- ✧ *Attest auditors must verify the system and controls to support the reliance they are placing thereon in expressing an opinion on the financial statements. Management auditors must verify the system and controls to ensure the reliability of their report to management on control adequacy and to ensure the appropriateness of their recommendations for improvements. The scope of the verification required to meet the differing objectives of these two sets of auditors will vary correspondingly.*

However, irrespective as to what the objective of the various auditors might be, it remains a fact that the analysis of systems that include data processing activities is a mutual concern to all types of auditors. (Mair, Wood & Davis, 1986:3; Porter & Perry, 1984:8)

Brown (1991:1) add that the internal controls of systems need to be evaluated to predict the reliability of systems, an approach that is commonly known as compliance auditing, which is the premise of the discussions in this chapter.

#### **4.2.1 Internal controls defined**

The American Institute for Certified Public Accountants in their Statement on Auditing Standards no 1 (SAS No. 1) subdivides internal control into



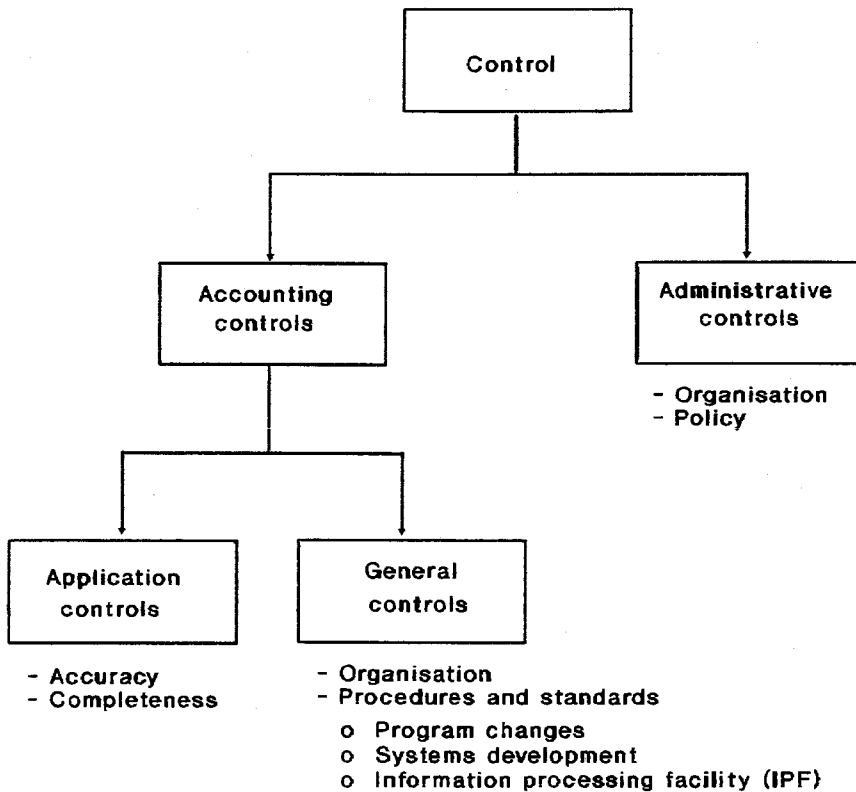
administrative and accounting control (AICPA 1972:para. 27&28) and explain:

- administrative control includes, but is not limited to, the plan of organisation and the procedures and records that are concerned with the decision processes leading to management's function directly associated with the responsibility of achieving the objectives of the organisation and is the starting point for establishing accounting control of transactions
- accounting control comprises the plan of organisation and the procedures and records that are concerned with safeguarding of assets and the reliability of financial records and consequently are designed to provide reasonable assurance that:
  - transactions are executed in accordance with management's general or specific authorisation
  - transactions are recorded as necessary (1) to permit preparation of financial statements in conformity with generally accepted accounting principles or any other criteria applicable to such statements and (2) to maintain accountability for assets
  - access to assets is permitted only in accordance with management's authorisation
  - the recorded accountability for assets is compared with the existing assets at reasonable intervals and appropriate action is taken with respect to any differences

In Statements on Auditing Standards No. 3. (The Effects of EDP on the Auditor's Study and Evaluation of Internal Control: SAS No. 3) accounting controls are further broken down into general and application controls and explained as follows (AICPA 1974:para. 7&8):

- general controls comprise (a) the plan of organisation and operation of the EDP activity, (b) the procedures for documenting, reviewing, testing and approving systems or programs and changes thereto, (c) controls built into the equipment by the manufacturer (commonly referred to as "hardware controls"), (d) controls over access to equipment and data files, and (e) other data and procedural controls affecting overall EDP operations
- application controls relate to specific tasks performed by EDP. Their function is to provide reasonable assurance that the recording, processing, and reporting of data are properly performed. There is considerable choice in the particular procedures and records used to effect application controls. Application controls often are categorised as "input", "processing" and "output" controls

Porter and Perry (1984:139) in combining the discussions of controls from SAS No. 1 and SAS No. 3 explain that application controls relate primarily to the accuracy and completeness of data within a specific application in an organisation. They point out that general controls, in turn, provide the standards and guidelines under which employees function in their work governing the functions of developing, maintaining and operating systems to process data. On the other hand, administrative controls are controls related to the authorisation procedures that determine the legitimacy of a function to perform specific acts and the authorisation process governing origination of transactions. The combined view of these controls within an EDP environment is depicted in Figure 4.1.

**FIGURE 4.1      CONTROLS IN AN EDP ENVIRONMENT**

Source: Porter W.T. & Perry W.E. 1984. EDP Controls and Auditing. Boston:Kent. (page 140).

#### **4.2.2      The external auditor's responsibility regarding internal control**

The independent auditor's basic objective is to express an opinion on whether or not the client's financial statements fairly present the financial position and results of operations in conformity with consistently applied generally accepted accounting principles (GAAP). In this regard the auditor's investigation must be conducted under generally accepted auditing standards (GAAS). The use of computer or even data base systems to produce the financial and operational statements of the organisation does not change this basic objective and the introduction of highly sophisticated recording media (tape or disk) and processing device (computer) does not

reduce the auditor's professional standards (Techavichit 1979:73; SAICA 1989:para. .06).

#### 4.2.2.1 Auditor responsibility for the evaluation of internal control according to general auditing standards issued by the South African Institute of Chartered Accountants

In February 1976, the South African Institute of Chartered Accounts (SAICA 1986a) issued a statement on Generally Accepted Auditing Standards, No. AU010. The Statement lists a number of general standards, standards of fieldwork and reporting standards that must be maintained during an examination leading to the expression of an opinion on financial statements. Audit standards are required to serve as procedural guidelines to be adhered to in each audit engagement as Davis (1968:23) points out:

*\* Auditing standards ... are independent of the personnel or machines used to process and maintain the accounting and financial records. The auditing standards must be broadly based in order to have applicability to a wide range of audit situations. However, they still relate specifically to each audit examination since they refer to an acceptable level of quality which must be maintained by the auditor in the selection and application of appropriate auditing procedures. Hence, auditing standards are also guides to procedures.*

The standards of fieldwork have a major impact on work performed by external auditors in evaluating a client's system of internal control as it states that the auditor has to study and evaluate appropriate internal controls as a basis for reliance thereon and for the design and determination of his audit procedures.

Evaluating a client's system of internal control forms a very crucial part of the external auditor's responsibility. In this regard the South African Institute of Chartered Accountants issued a standard of fieldwork entitled Accounting Systems and Internal Control, No. AU230 dealing with internal control evaluations.

Auditing statement AU230 requires the auditor to obtain an understanding of his client's accounting system and related internal controls. The auditor's design of auditing procedures will be based on this understanding. Controls should be evaluated in terms of their reliability, whether they are operating as designed, and their ability to prevent or detect errors. Each control, in turn, should be considered individually during the evaluation.

When the auditor decides to rely on a particular control he should consider whether the control (under consideration) depends on other controls and under which circumstances the other controls also need to be studied and evaluated (SAICA 1986:para. .18).

The reliance on internal controls is based on:

- (only) the controls taken into consideration – there may be controls which the auditor does not study and evaluate because they may be irrelevant or inadequate (SAICA 1986:para. .11)
- an acceptance by the auditor that the effectiveness of internal controls are limited due to management's cost effectiveness requirement, the potential for collusion and human error, or management's action in overriding controls (SAICA 1986:para. .11-.13)

#### 4.2.2.2 Auditor responsibility for the evaluation of internal control in a computer environment according to the audit and accounting guide issued by the South African Institute of Chartered Accountants

South African Institute of Chartered Accountants issued in June 1989 an auditing and accounting guide named Auditing in a Computer Environment. The guide is supplementary to the statements on generally accepted auditing standards and does not have the authority of these statements. Nevertheless the auditor should have good cause to deviate from these guidelines, or he may be held liable if his opinion is in doubt.

Audit objectives and the applicability of generally accepted auditing standards do not change when financial information is processed by the computer but the audit process may, however, be affected (SAICA 1989:para. .06).

Where an entity's accounting systems are wholly or partially computerised it will be necessary for the auditor to obtain an understanding of the computer environment. The guide stipulates (SAICA 1989:para. .37-.39):

*... This understanding would comprise part of the auditor's assessment of the adequacy of the accounting systems as a basis for the preparation of financial information. A preliminary understanding of the computer environment and the computerised application's would be necessary to assist the auditor in designing auditing procedures.*

The guide gives recognition to advanced EDP applications which, according to the guide, are characterised by one or more of the following characteristics (SAICA 1989:para. .56):

- absence of input documentation
- authorisation of input controlled by the computer
- online data input
- real-time update
- a single transaction updating multiple files
- extensive use of programmed controls
- application controls not evidenced by computer output
- absence of management or audit trail

The controls needed in advanced EDP systems are more complex than in the case of other computer systems. Two categories of controls are discussed which include the following control approaches:

- Application controls

Programmed controls such as computer sequencing and matching; edit, validation and reasonableness checks; access authorisation; exception

reporting on unusual, inconsistent or material items; monitoring uncleared items; run to run totals (SAICA 1989:para. .61-.66)

- Environmental or General controls

Controls in respect of systems development, implementation and program change; access to programs and data; computer operation; physical access, security, backup and recovery; system software, organisation and administration of computer function (SAICA 1989:para .74)

Since programmed controls are often not evidenced by output from the computer, reliance on environmental controls increases. With these controls, two types of controls directly affect the audit because of their impact on the organisation's financial information, namely (SAICA 1989:para. .75-.78):

- system development, implementation and program change control
- controls over accessing programs and data

#### 4.2.2.3 Auditor responsibility for the evaluation of internal control in terms of supplementary auditing standards relating to EDP environments issued by the American Institute of Certified Public Accountants

In the absence of any specific standard by the South African Institute of Chartered Accountants relating to auditing in EDP environments the discussions will concentrate on specific auditing standards issued by the American Institute of Certified Public Accountants (AICPA).

### STATEMENT ON AUDITING STANDARDS NO. 3

Statement on Auditing Standards No.3 (SAS No.3) named The Effects of EDP on the Auditors Study and Evaluation of Internal Control was issued during 1974 to provide a framework for the development of further guidance concerning auditing procedures in examining financial statements of entities that use EDP in accounting applications. This statement describes the

effects of EDP utilisation on the various characteristics of accounting control and on the auditor's study and evaluation thereof (AICPA 1974:para.5).

SAS No.3 divides the auditor's study and evaluation of accounting (general and application) controls into a preliminary and a completion phase of the review.

During the preliminary phase the auditor identifies and understands the structure of EDP accounting controls as related to a particular client's system of internal control. Essentially, SAS No. 3 implies that the following requirements should be met in terms of the preliminary review phase (Davis & Perry 1982:44):

- the basic *flow of transactions* through each significant EDP application and the extent of EDP use in the application must be understood
- the basic *control structure*, including both manual and computerised portions, must be understood
- the *content of EDP files* and the relevance of the data to the financial statements and the audit must be understood

The auditor's preliminary understanding ordinarily is obtained by inquiry, but it may also be obtained by observing client personnel and reviewing documentation. At the end of the preliminary phase of the review, the auditor must decide on the adequacy of the internal accounting controls. (AICPA 1974:para. 25, 26)

If the auditor decides to complete the review, audit procedures would be developed to determine how individual controls function, their relationship to other controls, and, through compliance testing, whether they are functioning as intended. In this regard SAS No. 3 states (AICPA 1974:para. 27):



*The purpose of tests of compliance is to provide reasonable assurance that accounting control procedures are being applied as prescribed. Tests of compliance are concerned primarily with the questions: (a) Were the necessary procedures performed? (b) How were they performed? (c) By whom were they performed?*

#### STATEMENT ON AUDITING STANDARDS NO. 48

In July 1984, the AICPA issued Statement on Auditing Standards No. 48 entitled The Effects of Computer Processing on the Examination of Financial Statements. This Statement is the external audit profession's first attempt to develop new auditing guidelines for complex EDP systems. SAS No. 48 supersedes sections of SAS No. 1 and SAS No. 3, integrating pertinent existing guidelines with proposed new guidelines. SAS No. 48 apply auditing standards to the computer environment in the following areas (Reilly 1986:6):

- planning and supervision
- analytical review
- study and evaluation of internal control
- evidential matter

In commenting on SAS No. 48, Reilly (1986:7) states that in accordance with the new standard, as part of the planning process, the auditor should consider:

- the extent to which the computer is used in each significant accounting application
- the complexity of the entity's computer operations, including the use of an outside service center
- the organisational structure of the computer processing activities
- the availability of data

The specific provisions related to the study and evaluation of internal control are the major portion of SAS No. 48. The methods which the organisation uses to process significant accounting applications, may influence the control procedures designed to achieve the objectives of internal accounting control. It is clear that the auditor should consider the following factors in his study of EDP internal control (Reilly 1986:7):

- existence of transaction processing trails
- uniform processing of transactions
- adequate segregation of duties
- potential for errors and irregularities
- potential for increased management supervision
- initiation or subsequent execution of transactions by the computer
- dependence of other controls on EDP controls.

#### 4.2.2.4 Auditor responsibility for the evaluation of internal control in terms of supplementary auditing guidelines relating to EDP environments issued by the International Federation of Accountants

The International Auditing Practices Committee (IAPC) of the International Federation of Accountants (IFA) has issued International Guideline 20 (1985) named The Effects of an EDP Environment on the Study and Evaluation of the Accounting System and Related Internal Controls and Supplements 1, 2 and 3 for EDP systems. The supplements address the specific audit requirements of *Stand-alone Micro Computers (1987)*, *Online Computer Systems (1989)* and *Data Base Systems (1989)* respectively. The supplements provide specific details as to how the design of accounting systems and related internal controls is affected by the various EDP

environments. The guideline classifies controls into general EDP controls (overall controls) and EDP application controls (the specific controls over the accounting applications). (Fink 1991:13)

In terms of the objectives of this study only supplement 3 (addressing data base systems) will be discussed. Supplement 3 identifies the following internal controls as being particularly important due to the specific characteristics of the data base environment (Fink 1991:16):

- General EDP controls
  - a standard approach for the development and maintenance of application programs
  - clear and definite assigned responsibility for data ownership
  - password security over access to the data base
  - segregation of duties with regard to the data design
  - administration and operation of the data base
- EDP application controls (which rely on general EDP controls as described)
  - Data base management system specific: data independence from application programs; data interrelationship safeguards; use of data dictionary
  - Data administration: correct data base structure definitions; rules for data integrity, completeness and access; coordination of computer operations related to the data base; monitoring system performance; providing administrative support

#### **4.2.3 The internal auditor's responsibility for internal control in manual and automated environments**

The evaluation of the system of internal controls also forms part of the responsibilities of internal auditors. As business organisations have grown over the years management sought reports on what has been happening and why. Internal auditing was developed in order to report to management.

The Institute of Internal Auditors (IIA), during 1978, adopted standards for evaluating the activities of internal audit departments known as the Standards for the Professional Practice of Internal Auditing. These standards prescribe the following (IIA 1978:5-7):

- internal auditors should be independent of the activities which they audit – Independence: Standard 100
- internal audits should be performed with proficiency and professional care – Professional Proficiency: Standard 200
- the scope of the internal audit should encompass the examination and evaluation of the adequacy and effectiveness of the organisation's system of internal control and the quality of performance in carrying out assigned responsibilities – Scope of Work: Standard 300
- audit work should include planning the audit, examining and evaluating information, communicating results, and following up – Performance of Audit Work: Standard 400
- the director of internal auditing should manage the internal auditing department properly – Management of the Internal Auditing Department: Standard 500

Standard 300, Scope of Work, deals specifically with the internal auditor's responsibility pertaining to evaluating an organisation's system of internal

control. Paragraph .02 of this standard further adds that the purpose of the review of the system of internal control is to ascertain whether the system established provides reasonable assurance that the organisations objectives and goals will be met efficiently and economically (IIA 1978:17).

In addition to the auditing standards issued by the Institute of Internal Auditors (IIA) the institute also adopted two prominent research reports that give direction to internal audit work.

During 1975 the IIA, with a grant from IBM, commissioned a research study with Stanford Research Institute entitled Systems Auditability and Control Study: Data Processing Auditing Practices Report which covered a two-year period involving 45 organisations in the United States, Canada, Europe and Japan as well as a mail survey of 1,500 organisations. The study explored the state of the art in the audit and control of computerised information and data processing systems to identify current trends in auditing and control. It documented specific audit and control techniques in use at that time which still offers some practical value to internal auditors. Some of the main conclusions are (IIA 1977:4-7):

- the primary responsibility for overall internal control resides with top management, while the operational responsibility for the accuracy and completeness of computer-based information systems should reside with the users
- there is a need for improved controls because inadequate attention has been given to the importance of internal controls in the data processing environment. Control objectives should also be identified during the systems development process and should be addressed as separate development requirements
- internal auditors must participate in the system development process in order to ensure that appropriate audit and control features are designed into new computer-based information systems since an evaluation of the

adequacy of controls after a system has been installed will determine any weaknesses too late in the development process

- verification of controls must take place both before and after installation of computer-based information systems
- the need exists for greater internal audit involvement in the data processing environment as a result of the growth in complexity and use of computer-based information systems
- many organisations are not adequately evaluating their audit and control functions in the data processing environment. Top management should initiate a periodic assessment of its audit and control programs

Most of the conclusions directly affect an organisation's system of internal control. The study further emphasised that new data processing control techniques and internal auditing approaches are needed to ensure accuracy, completeness, timeliness, and security of computer application programs.

A more recent study conducted by the Institute of Internal Auditors Research Foundation Systems Auditability and Control Report: Modules 1-12 pertaining to the internal auditor's evaluation of internal control confirms most of the findings of the 1975 research but also gives new direction to internal auditors. The main findings relating to control objectives and risks are (IIARF 1991a:2-41; 2-42):

- ensuring that an adequate system of internal control is in place in the organisation is the responsibility of management
- the system of internal control includes the processes, functions, activities, subsystems, procedures, and organisation of human resources that provide reasonable assurance that the goals and objectives of the organisation are achieved and assure that risk is reduced to an acceptable level

- the use of information technology poses risks to information and process integrity that must be mitigated through the implementation of cost effective controls
- internal auditors should review the system of internal controls to verify their presence and continuing operation. In order to fulfil this responsibility, auditors should understand the technology and the resulting risk and control implications
- advancements in information technology and the reliance of organisations on information systems continue to provide significant challenges both to management and internal auditors

#### **4.2.4 Modern auditing practice pertaining to audit involvement during systems development**

It was explained in sections 4.2.2 and 4.2.3 that auditing standards pertaining to internal and external auditors require an appropriately organised study and evaluation of the internal controls of a particular system under review.

In encountering large, complex and highly integrated systems with development hours that can easily total up to 100 man years it becomes difficult for the auditor to conduct an audit of the controls after system implementation. It therefore makes a lot of sense for the auditor to conduct his audit of controls during the various stages of development. For a large system it is prohibitively expensive to make any significant modifications after its implementation.

Davis and Perry (1982:32) explain:

*From a historical perspective, the traditional internal control review often provides too little too late. It takes place after a new system is implemented, if not much later. It often results in a long series of control-oriented recommendations that are too costly to implement*

*once the system is up and running.*

Nasuti (1987:1) supports this view by propagating that the audit process should commence before a system is up and running and auditors should therefore be involved while the system is being designed and developed. He maintains that auditors should be consulted during the initial phases of systems development regarding control decisions and recommends that audit reviews of the system be conducted throughout the development life cycle.

Sardinas, et al. (1981:8), in line with Nasuti, confirm this approach in the following way:

*Auditors, both internal and external, must become involved in the initial design and testing of computer-based systems to ensure that they contain sound controls and are, in fact, auditable. This is an extremely important point. Changing a system after it is implemented often is very costly and invariably is strongly resisted. Auditors should be involved from the very beginning and should be the major advocates of proper controls.*

In support of a participatory approach the Canadian Institute of Chartered Accountants (CICA), in their auditing guideline entitled Auditing in an Electronic Data Processing Environment, stipulates that it may be appropriate for the auditor to commence a preliminary evaluation of controls within a new system during the system's development process (Wingate 1981:76).

In a more clearly worded standard, the Government Accounting Office (GAO) of the United States in their supplementary standards of 1981, states that the auditor shall actively participate in reviewing the design and development of new data processing systems or applications, and significant modifications thereto, as a normal part of the audit function. It is also stipulated that the auditor shall review general controls in data processing systems to determine whether they have been designed according to management direction and known legal requirements and are operating effectively to provide reliability



of, and security over, the data being processed. The auditor shall further review application controls of installed data processing applications on which the auditor is relying to assess their reliability in processing data promptly, accurately and completely. (Wysong 1983:94). This particular standard is only applicable to government auditors in the United States of America.

The Systems Auditability and Control (SAC) research project conducted by Stanford Research Institute for the IIA (1977) highlighted the fact that internal auditors should participate in the systems development process to ensure that the necessary audit and control features are built into new computer-based information systems (Perry 1985:3).

The Institute of Internal Auditors Research Foundation's SAC report of 1991 (in line with their previous report of 1977) states that the internal auditor should be involved in the systems development process (IIARF 1991b:5-77).

The view is expressed that the auditor should be aware of and consider new computer systems during the audit planning process. The auditor should be concerned with both general and application controls as well as the overall control structure including the manual and computerised portions of the total system as outlined in the auditing standards, guidelines and research reports.

It is clear from the discussions of the various auditing standards, supplementary standards and guidelines, that the basic objective of the system of internal control within the data base environment remain, i.e., to provide reasonable, but not absolute assurance, that assets are safeguarded and that (financial) records provide a reliable basis for the preparation of financial and management statements and reports.

In order to ensure the integrity of an organisation's information, auditors should evaluate controls over data, transactions and programs. The responsibility of the auditor is to identify the relevant and appropriate controls and to test such controls for evidence of ongoing compliance.

The view is further expressed that the auditor should be involved in systems development right from initial planning phase to implementation to be in a position to conduct a meaningful evaluation of the internal controls especially in terms of large and complex data base systems which are the premise of this study.

#### **4.2.5 Internal / External auditor cooperation in participating in systems development**

Chen (1983:75) expresses the view that on a practical level, few external auditors have the time or flexibility needed for comprehensive systems development involvement as part of their pure attest audit engagement. To overcome this problem, he believes, the external auditor should use the skill and flexibility of the internal auditor in systems development participation.

Watne and Turney (1990:173) maintain that when internal auditors are actively involved in the systems development process, they themselves become a general control which contributes to the reliability of systems.

Chen (1983:224) explains that the internal audit profession has made big strides in EDP auditing over the last decade. He substantiates his reasoning by saying:

*As corporations computerized and moved to more advanced systems, so too, did the demands made on internal auditors to understand and audit these systems increase. It was not uncommon for entire internal audit departments to be replaced if the existing staff could not cope with the pace of new technology. Many large corporations today have a separate EDP audit function, staffed by highly trained and knowledgeable specialists.*

In terms of the dilemma that external auditors are facing Chen (1983:223) further argues that due to the nature of public auditing practice, the external auditor serves a large number of clients, each of whom uses a different EDP system which makes it extremely difficult for computer audit

staff to specialise and become highly knowledgeable about certain advanced systems. Not to be forgotten, of course, is the big expense involved in training, updating and retaining technical staff, he further argues.

No South African statement on auditing deals directly with the issue of internal/external audit cooperation. South Africa, however, is a member of the International Federation of Accountants and is thus, in the absence of a specific statement dealing with this topic, implicitly bound by that body's guidelines. In this respect, circular SAI 2/82 of the South African Institute of Chartered Accountants titled The Relationship Between Statements Issued in South Africa and Guideline Statements Issued by the International Federation of Accountants indicates, in paragraphs 7.4 and 7.5, that if no South African statement on auditing exists then the international statement dealing with the particular topic will be applicable (Taylor 1983:26).

The International Auditing Guideline (IAG) 10 entitled: Using the work of an Internal Auditor imply that reliance on the internal audit function is feasible where the following procedures have been conducted by the external auditor concerning the evaluation of internal audit work performed (Taylor 1983:72), namely:

- examination of the work by the internal auditor
- examination of work similar to that conducted by the internal auditor
- observation of the internal auditor's procedures

Chambers and Court (1991:32) comment that the extent to which the external auditor will rely on internal audit work is further related to aspects such as:

- the degree of internal audit independence
- the materiality levels to which the internal auditor works

- the level of professionalism of the internal auditor and quality of work

It is believed, in terms of the scope of this study, that internal and external auditors should cooperate, in the evaluation of controls pertaining to systems under development, in a practical manner as implied by International Auditing Guideline (IAG) 10. In this regard, the approach propagated by Chen in terms of the external auditor / internal auditor cooperation in system development, is supported since the internal auditor would be an ideal person to carry out this evaluation and report on the findings to the external auditor. The internal auditor is normally also in a better position to conduct the necessary thorough study and evaluation of the system of control pertaining to the system under development.

The roles of and work performed by the internal auditor should complement the work of the external auditor. In this way duplication of effort is avoided and the external auditor can report on the financial statements with a greater degree of certainty. In addition management can feel more secure about the strength of internal control systems as well as the accuracy and fairness of financial reporting (Gilhooley 1986:9).

Davis and Perry (1982:21) supports this view by saying that auditing effort should be considered in its "totality" and not split between internal and external auditing - an emphasis toward much more coordination of the total audit effort.

#### **4.2.6 The auditor's objectives for participating in the systems development process**

Weber (1982:99) is of the opinion that the auditor participates in the system development process to ensure that controls are built into the system to safeguard assets, to ensure data integrity and to achieve system effectiveness and efficiency. The auditor collects evidence primarily by observing the activities of the other members of the development team.

Gallegos and Dow (1987:3) support this view but highlight the fact that the view of the auditor should mainly be that of reviewing and evaluating the adequacy of controls during the systems development process and of commenting on the auditability of the system being developed.

Nasuti (1987:1) emphasises that one must remember that it remains the responsibility of management to ensure that information systems have a satisfactory level of internal control. The role of the auditor, in turn, is therefore one of purely assisting management in accomplishing this objective. The role of the auditor, he maintains, is to evaluate the adequacy and effectiveness of the controls, to test compliance with these controls and then to report back to management his opinions and recommendations.

In terms of this study the objectives of the auditor for systems development participation will be restricted to that of gaining an early understanding of the overall development process in order to evaluate and test the adequacy and effectiveness of controls in predicting the reliability of information systems.

#### **4.2.7 The scope of the auditor's participation in the systems development process**

According to Nasuti (1987:3) the scope of the auditor's involvement in the systems development process should be that of determining the level of internal control per life cycle phase and advising the systems development team on matters pertaining to internal control issues.

Within the ambit of this scope Nasuti (1987:3) believes that the work should also include:

- an appraisal of the systems development methodology to determine whether it is in accordance with the organisation's operating practices
- the compilation of reports and recommendations to management

### **4.3 ENSURING THE RELIABILITY OF AN ORGANISATION'S INFORMATION SYSTEMS WITH SPECIFIC REFERENCE TO DATA INTEGRITY AND DATA MANAGEMENT**

It can be said that the management of information has become a key discipline within every organisation. Information forms the basis for decision making, and the quality of decisions is directly related to the accuracy of the information provided to decision makers. For information to be reliable it must be appropriate for the use to which it is to be put and available when required, explains the British Computer Society (1989:36). The reliability of information depends on the integrity of the processes which transforms "raw" data into information, as well as the integrity of the raw data itself (IIARF 1991b:5-13, 5-14).

In terms of this study the discussions will concentrate on the data component of information systems (explained in Section 4.1). Within the context of this qualification Moss (1983:1) asserts that an organisation's information needs can only be met when the underlying data is readily available and accurate.

As organisations are becoming more dependent on information, ensuring the appropriate level of control over data and its underlying data base technology has become more important and, in many cases, crucial for survival.

#### **4.3.1 Ensuring data integrity in a data base environment**

##### **4.3.1.1 Data integrity described**

Data integrity means different things to different people and will probably continue to do so for some time to come in spite of considerable effort to establish a consensus definition. For the purpose of this study system integrity will be defined as the measures taken to ensure the correctness of information (Butler Cox Foundation 1988:50).

The term "integrity" as it relates to data will be described, within the context of this study, as the measures taken to ensure the correctness (i.e. data is exactly what it purports to be) of the data in the data base.

The Butler Cox Foundation (1988:50) explains:

*These measures ensure that the database structure always reflects the structure of the data model used to design it, and that the data it contains conforms to the descriptions in the data dictionary.*

Percy (1986:123) argues that the "integrity" of data can only be ensured when attention is given to the "quality" aspect of data throughout all life cycle phases of systems. He explains:

*First, the relevant data have to be identified. They must be designed properly so they represent business facts accurately. Next, the processes that will work on those data must be specified precisely and unambiguously. The implementation of those processes must be verifiable and correct so that information pollution does not occur. Finally, the data presented must be applied wisely.*

Thompson (1989:365) extends this argument by saying that any discussion on integrity of data is really a discussion on "data modelling". He believes that there are different types of integrity aspects at the different levels of the system. The issue of logical integrity is, he explains, that of deriving at a *physical representation of the logical data schema* (data model) that is correct and manageable in terms of its data content.

Logical data integrity in terms of the scope of this study will be defined as the measures taken to ensure that the physical data base structure always reflects the structure and meaning of the data model used to design it. (Butler Cox Foundation 1988:50; Thompson 1989:388).

#### 4.3.1.2 The importance of data integrity within a data base environment

Durell (1985:1) maintains that an organisation's information needs can only be met when quality data is available to support those needs. He explains that quality depends on the integrity of the data – that is, its accuracy, reliability, timeliness, consistency and standardisation.

Data integrity is therefore a fundamental concept. It is a state which implies that data has certain attributes of completeness, soundness and purity. Without data integrity being maintained, an organisation no longer has a true representation of itself or of real world events. Maintaining data integrity can only be achieved at a cost and the benefits obtained should exceed the costs of the control procedures required.

In order for users to feel assured of the necessary degree of accuracy, formal procedures must exist for the maintenance of data integrity at a satisfactory level. Controls that allow management to make decisions based on bona fide information processed in a controlled environment are critical. Controls should be designed and implemented with the objective of ensuring integrity of data and its underlying data base technology.

Weber (1982:8) explains that the value of the informational content of the data for individual decision makers, and the extent to which the data item is shared among decision makers are the two major factors that affect the value of a data item to an organisation.

Perry (1987:1) points out that a data base environment, in which a single application or a series of closely related applications is the only user of the data base, is not much more complex than the traditional file environment. This kind of data base environment in its simplest form may not pose more integrity problems than a non-data base environment.

Techavichit (1979:16, 17) expresses the view that in terms of a more sophisticated data base environment with the specific objective of shared



data, typified by a number of users making use of the common data base, necessitates the establishment of data integrity as an important control objective of the organisation and its management.

Weber (1982:8) points out that corruption of data integrity affects not only one user, but multiple users in an environment where data is shared. The value of a data item, he believes, is the aggregate function of the value of the data item to the individual users of the data item.

The view is expressed that data integrity, within the overall context of the data base environment, can only be maintained when appropriate controls are in place to ensure that:

- data is identified correctly in respect of the organisation's information requirements
- proper design of data takes place to represent business facts accurately
- processes that will work on those data to be specified precisely and unambiguously
- the implementation of those processes to be verifiable and correct so that information pollution does not occur
- in the last instance, that whenever data (information) is presented it has to be applied wisely

#### **4.3.2 Primary risk related to information system development within the context of an IRM approach to systems development**

It is believed that within the context of an IRM emphasis to systems development the primary risk is that of information systems which are developed and do not meet the business needs.

A recent survey, in support of this view, listed the highest information related risk as that of systems that are incompatible with organisational objectives which result in business needs not being met (IIARF 1991b:5-3). The survey's findings is further in line with the risk of inadequately defined systems which result in competitive disadvantages highlighted by Mair, Wood & Davis (1986:210) and Biggs, Birks & Atkins (1980:4).

The data-related risks pertaining to the primary risk of information systems that are incompatible with organisational objectives are (IIARF 1991b:5-40):

- Ineffective design

The risk relates to the design of the data bases that do not meet the organisation's requirements due to insufficient or inefficient specifications during the planning and initial development process. Improperly designed data can result in the unavailability of vital information.

- Invalid or unidentified relationships

Invalid relationships present a false or inaccurate picture of data usually by linking together unrelated data items. Unidentified relationships, in turn, can result in the inability to bring together related items.

- Inconsistent data

Inconsistent data (data derived from different assumptions and used as if the same) can result in poor decisions based on:

- inconsistent cut-offs
- meaningless comparisons of data
- invalid consolidation of data

- Lack of clarity of definitions

Missing or unclear data definitions can result in information that is not linked to all other relevant information in the system.

- Lack of appropriate ownership

The inability to determine proper ownership may result in insufficient data definition.

It is believed that, without closely linking the systems development and data management practices with the organisational objectives, the foundation of applications will be based on a dubious set of requirements and achieving control will be difficult. Systems that are not aligned to organisational strategy will not satisfy business or user requirements. The compatibility of information systems and the commensurable data management practices with organisational objectives can mitigate the risk of systems not meeting user needs.

The Institute of Internal Auditors expresses the view that the most important management challenges of the 1990's will be to integrate the planning, design and implementation of complex application systems with the strategy of the organisation to ensure that user needs are met (IIARF 1991c:1-1, 1-2).

#### 4.3.3 The role of data administration in ensuring the integrity of an organisation's data

Porter and Perry (1984:180) maintain that the aspect of data integrity has a pervasive effect on the data management functions. They argue that although total integrity protection is never completely achievable there are, however, several integrity controls that can be invoked by data management to achieve an effective level of integrity. They (1984:180-181) explain that these controls range from *the preventive measures of data definition control through the assurance measure of backup and recovery*.

The most prominent procedures to be utilised to achieve and maintain integrity are (Porter & Perry 1984:181; IIARF 1991a:2-21):

- Control procedures related to logical integrity (data-related)

- control over the data definition which ensures that data conform to the definition and that conformance is maintained
- Control procedures related to physical security (data base-process-related)
  - access control that protects the data from access by unauthorised persons or for unauthorised purposes
  - maintenance of an audit trail to demonstrate that the confidentiality of data is intact
  - update controls that ensure the user has the appropriate authorisation to change data values
  - concurrency controls that provide integrity by controlling programs that perform the update function

The discussions in terms of this study will only concentrate on the logical data integrity aspects of systems. In this regard Thompson (1989:363,388) explains that logical integrity is only concerned with ensuring that data is correct as it relates to a common understanding and meaning among the individuals using the data.

It therefore follows that ensuring logical integrity of data should be the responsibility of the organisation's DA as this function is responsible for (refer Section 3.4.1):

- determining the information needs of the organisation
- creating and maintaining the corporate data models, ensuring that they are as stable as possible
- obtaining agreement among users about definitions and format of data

items

- ensuring that system builders conform to the data models
- resolving conflicts about incompatible representations of data

The DA liaises with end-users, senior management, systems development and data base teams to define each data element – its size, format, usage and other pertinent meta-data which is then properly documented in a data dictionary (Section 3.4.1.3).

The view is expressed that a DA function is a keystone to ensuring the logical integrity of the organisation's data.

#### **4.4 THE IMPACT OF A FORMAL DATA MANAGEMENT APPROACH WITHIN THE CONTEXT OF A DATA BASE ENVIRONMENT ON THE ORGANISATIONAL CONTROL STRUCTURE AND AUDITING**

The discussions of the previous chapter and sections have indicated that systems development within the context of a data base environment are radically different from traditional file – oriented systems. It was also illustrated in Section 4.3.3 that a formal data management function plays a vital role in ensuring the logical integrity of data. It is therefore reasonable to expect that a formal data management function can enhance the system of internal controls and that some of the internal controls used effectively within the file environment may not be effective or feasible within the data base environment.

##### **4.4.1 Control objectives and risks**

The control of an organisation's information systems is a very complex topic. Some of the complexity is caused by the wide range of controls required to ensure the successful operation of an organisation and its systems. On the one hand are such controls as top management's long-range planning and on

the other hand are accounting controls, such as programmed edit checks. Both of these types are important within the system of internal control because they, each in their own way, influence the reliability of an organisation's information systems, reports Roberts (1985:1.1).

Control objectives can be defined as statements of what the organisation wishes to accomplish by implementing controls. Any organisation strives to minimise its risks through the implementation of appropriate controls. It is management's responsibility to ensure that an appropriate system of internal controls is implemented for an organisation to meet its goals. (IIARF 1991a:2-19)

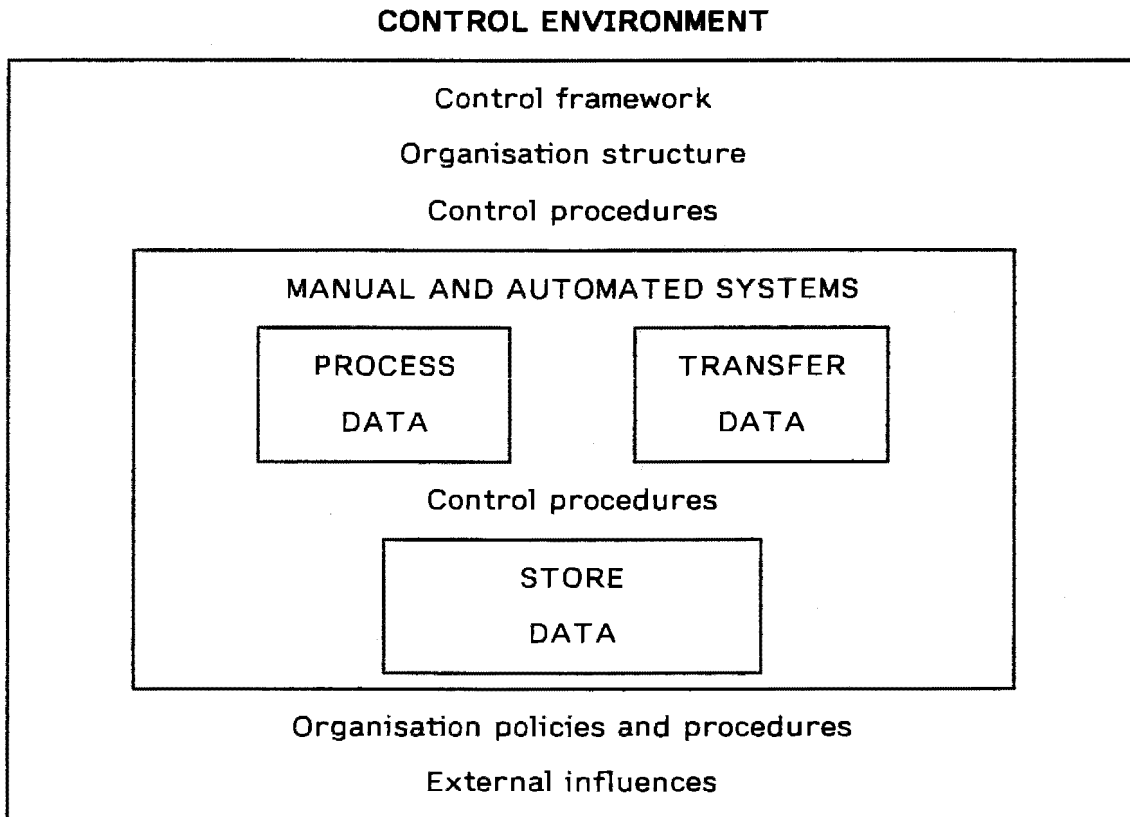
#### **4.4.2 Internal control**

The entire system of internal control can be defined as those processes, functions, activities, and people which are grouped together or specifically segregated to ensure the effective achievement of an organisation's objectives (IIARF 1991a:2-2).

Roberts (1985:1.3) explains that the system of internal control guides, regulates and directs an activity towards a desired result. The first element of the system of internal control is the control environment which influences and impacts the effective operation of all the internal controls. This environment establishes the conditions within which the systems and controls operate, and in that sense contributes positively or negatively on their reliability (IIARF 1991a:2-3).

The control environment depicted in Figure 4.2 includes the following:

- Organisational structure
- Control framework
- Organisation policies and procedures
- External influences
- Control procedures

**FIGURE 4.2 THE CONTROL ENVIRONMENT**

Source: Institute of Internal Auditors Research Foundation. 1991. Systems auditability and control report Module 2: Audit and control environment. Altamonte Springs, Florida (page 2-3).

The control framework covers such issues as segregation of incompatible duties, competence and integrity of staff, appropriate levels of authority and responsibility, accountability and supervision of staff planning mechanisms, policy, operating standards and other controls (IIARF 1991a:2-4). Within the ambit of this control environment manual and automated systems store, transfer and process data in the production of information as depicted in Figure 4.2.

Within the broader IRM context of this study the system of internal control can be viewed as all the controls implemented to ensure the achievement of an organisation's information requirements. In this regard Plagman (1982:8)

explains the system of internal control as those control procedures implemented to ensure integrity, namely accuracy, completeness and timeliness of information.

#### **4.4.3 The impact of formal data management on the organisational internal control structures**

Porter and Perry (1984:9) assert that the system of internal control normally has to be restructured because of the characteristics of a computer system. In this regard new controls have to be instituted which relate to the overall data processing environment, commonly known as general controls, as well as application controls which are part of and unique to a specific application system explains Wysong (1983:15 & 33).

Stoneham (1979:33) explains that the impact of data base technology on internal control vary, depending on the data base functions the organisation wishes to rely on and the complexity and extent of data base usage.

In terms of this study of addressing a more sophisticated data base environment with specific objective of shared data typified by a number of users making use of the common data base, many traditional internal controls, such as separation of duties and decentralised control of data files, which are effective in traditional systems are inadequate and ineffective (Techavichit 1979:14&15).

In a data base environment the methods of informational control and usage change from an application-oriented approach to an organisation-wide approach as explained in chapter 3. Within the ambit of traditional application systems, each application is a separate system with its own reporting and controls. In the context of a data base environment, however, many controls are now centralised and the data base is designed to serve the organisation more holistically. (Roberts, 1985:7.4).

Where data management activities in a non-data-base environment were



performed by an application programmer, they are performed by a data management team independent of the application project team in an integrated data base environment. In this regard the users and the application project team have to rely on the integrity of data managed by the data management team (Perry, 1987:1).

Perry (1986:7&8) claims that the introduction of data base technology requires a restructuring of the organisation by establishing a data management function which manages data in the same way as the financial manager and treasurer manage the cash resources of the organisation.

Ravindra (1987:1) points out that with the concept of viewing data as a corporate resource comes the need to manage data from an organisation-wide perspective. Weldon (1990:3), believes that in this regard centralised control is necessary for enacting standards, to promote data integrity for the organisation as a whole. Centralised data management should include data administration (DA) as well as data base administration (DBA) to ensure control, coordination, and integration (Ravindra 1987:1 & 2 and Perry 1986:3).

Holloway (1988:7) and Weldon (1990:3) support this view in saying that centralised control over data supports the information provision process by controlling and coordinating the definitions (meta-data, format and characteristics) and the usage of data. Coordination of activities is required to achieve an appropriate level of control and consistent interpretation of the meaning of data, claims the Joint Data Base Task Force (AICPA 1983:18).

The view is expressed that, in the context of an IRM approach to systems development, a centralised data management function forms an essential organisational control in ensuring information integrity and compatibility of information systems with organisational goals.

#### 4.4.4 The impact of formal data management on the auditing process in general

In auditing the emphasis has shifted from the evaluation and verification of processing results to more reliance on the evaluation and verification of controls designed and implemented to ensure the continuing accuracy and reliability of those results (Techavichit 1979:20; Chambers & Court 1991:39).

It is believed that the arrival of data base systems necessitates a further extension of the trend to put more reliance on the evaluation and verification of controls.

It is most important to note that the objectives of auditing, whether to meet the needs of shareholders, management or other parties, are not affected in any way by the manner in which information is processed, be it by computer or a manual system nor by the introduction of data base systems.

Wysong (1983:29) emphasises:

*Advanced EDP technology is having a substantial impact on the auditors' role. Audit requirements and standards have not changed. What has changed, however, is the auditors' approach to perform their traditional functions ...*

The view is expressed that, although the audit objectives and standards do not change in a data base environment, the introduction of data base technology has tended to increase the exposure that may result from poor control since the data base itself contains data from many sources that is shared by many users making use of the "common" data base. Techavichit (1979:72) is also of the opinion that *the trend toward database systems has heightened the pressure for increased care and more effective controls.*

The data base environment requires controls that span across individual applications at a high enough level to ensure control of data for the organisation as a whole – a concept known as the migration of *application*

*controls.* Many observers report that there is a continuing trend toward the migration of application controls from application systems to general controls in the general information system environment, as systems become more integrated and complex in the data base arena (IIARF 1991a:2-41; AICPA 1983:20 & Plagman, 1982:9). In this regard, the Institute of Internal Auditors (IIARF 1991a:2-42) argue that these general controls are more pervasive in that they affect multiple applications or business operations and their importance should be reconsidered by the auditor.

Following on this argument, the view is expressed, that the audit approach to auditing information systems in a data base environment should concentrate more on a control structure applied to the overall operating environment of information systems than the actual application level itself within the given scenario of some highly significant data base system developments. It is for this reason, it is believed, that the auditor must recognise the fact that more reliance can be placed on, and more audit satisfaction obtained, from effective general controls that relate to and condition the quality of the overall data base (Jenkins 1981:213).

As an effective data management function fulfils a key role in ensuring the integrity of the overall data base (explained in Section 4.3.3) it becomes an essential general control on which the auditor can rely in terms of his evaluation of the system of internal controls.

Formal data management also creates an opportunity for improving audit effectiveness since the data administrator will be responsible for setting the control philosophy. By working with the data administrator auditors will have the opportunity to help establish the control philosophy and develop an integrated approach to audit and control of the organisation's information systems (Rollier 1982:184; Borthick 1986:44)

Rollier (1982:184) is of the opinion that:

*Auditors should recommend or even demand that the data*

*administration function be established, with the skills and resources, time, and management backing it needs to accomplish its objectives.*

#### **4.5 THE IMPACT OF A FORMAL DATA MANAGEMENT APPROACH IN THE CONTEXT OF A DATA BASE ENVIRONMENT ON THE TRADITIONAL SYSTEMS DEVELOPMENT AUDIT PROCESS**

The objective of audit involvement in systems development was determined in Section 4.2.6 of this chapter as that of gaining an (early) understanding of the overall development process in order to evaluate and test the adequacy and effectiveness of controls. It was explained that the end objective is to predict the reliability of information systems.

The scope of the auditor's involvement in systems development was described as appraising the systems development methodology, determining the level of internal control per life cycle phase, advising the system development team on control issues and reporting on the audit findings.

The view is expressed that the above objective of the systems development audit and the overall scope of audit activities will remain unchanged in a data base environment where a formal data management approach in managing the organisation's data resources has been instituted. However, it is believed that the audit approach required in conducting the control evaluation per life cycle phase will differ considerably.

In determining the overall impact of a formal data management approach (in the context of a data base environment) on the traditional audit control evaluation process it is suggested, and necessary, to address the impact per life cycle phase. The discussions in terms of this study will not specifically address the activities of appraising the development methodology, the role of advising on control and reporting. A formal data management approach to systems development in the context of a data base environment will only be referred to as a data management approach in the discussions that follow.

#### 4.5.1 Data management and logical data integrity described in terms of the scope of this study

A data management emphasis to systems development in terms of this study implies a "data-driven" approach to systems development as explained in Chapter 3. It further implies that data analysis and design are conducted independently, but in conjunction with the analysis and design of application processes. This separation is crucial for the construction of a "generic data base" that is not bound by the constraints of individual applications, comments Gilhooey (1985:5).

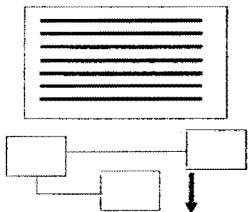
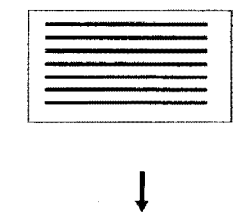
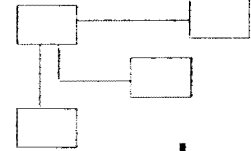
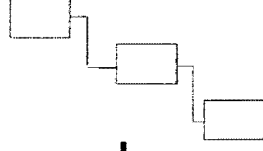
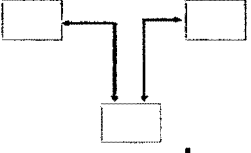
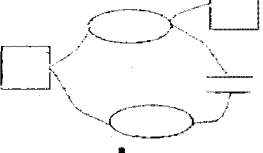
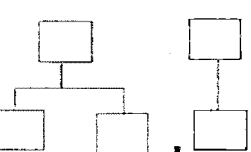
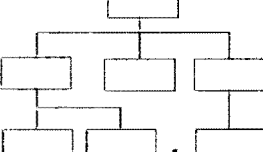
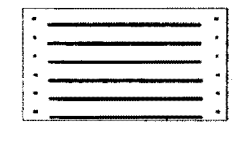
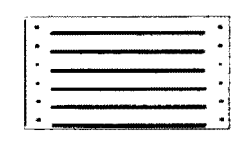
A data management emphasis also implies that a sound, logical framework or architecture for defining and controlling the interfaces and integration of the data and process components, for the orderly development of information systems, are in place. In this regard, the Institute of Internal Auditors (IIARF 1991b:5-20) claims that *the use of representations and descriptions provides the basis for the development of a descriptive framework (architecture) for data and applications as depicted in figure 4.3*. For every different type of description, there are different models that address the points of interest at the particular phase of the systems development as they (1991b:5-22) explain:

*There are significant differences in the representations during the progression of the system life cycle stages. Each representation is different from the others, and there is a refinement in terms of scope, system functionality, and purpose from one representation to the next. Each representation has its own distinct nature.*

The various models created during the development of systems, in terms of the scope of this study, are the result of a series of planned efforts and tasks outlined in the information engineering development life cycle methodology. These efforts and tasks culminate in the compilation of conceptual, logical and physical models that describe either the data or processes that act on the data. In terms of the data component of the information systems development process, the effort is aimed at the

construction of the physical data base based on the user requirements identified during the "data planning" phase and refined during the subsequent phases of conceptual and logical data model compilation.

**FIGURE 4.3 REPRESENTATION, DESCRIPTIONS AND MODELS SUPPORTING AN INFORMATION SYSTEMS ARCHITECTURE**

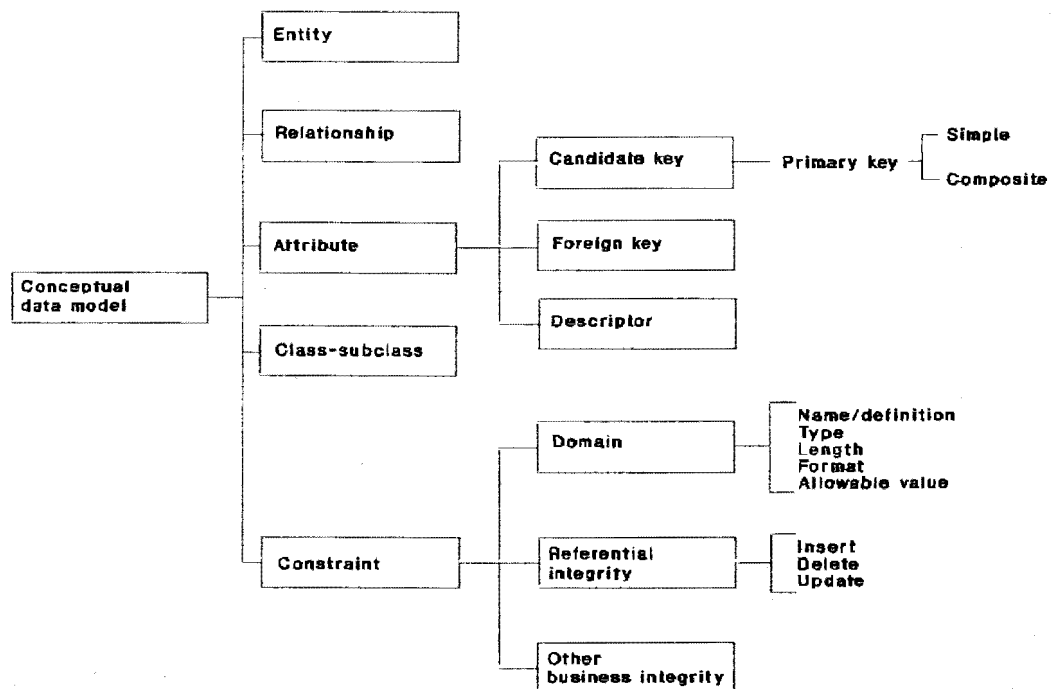
Description Representation	Data Description	Process Description
<b>Enterprise Model</b>  <b>(Ballpark View)</b>  NOTE: PHASE 1 PER FIGURE 3.4	e.g., List of Entities and high level entity model  	e.g., List of Processes the Business Performs  
<b>Conceptual Model</b>  <b>(Owner's View)</b>  NOTE: PHASES 2 & 3 PER FIGURE 3.4	e.g., Entity/relationship Diagram  	e.g., Functional Flow Diagram  
<b>Logical Model</b>  <b>(Designer's View)</b>  NOTE: PHASE 4 PER FIGURE 3.4	e.g., Data Model  	e.g., Data flow diagram  
<b>Physical Model</b>  <b>(Builder's View)</b>	e.g., Data Design  	e.g., Structure Chart  
<b>Detailed Description</b>  <b>(Out-of-Context View)</b>  NOTE: PHASE 5 PER FIGURE 3.4	e.g., Data description  	e.g., Program  

Source: Institute of Internal Auditors Research Foundation 1991. Systems auditability and control report - Module 5: Managing information and developing systems. Altamonte Springs, Florida (page 5-21).

Within the context of an IRM approach to systems development the data modelling effort must strive to ensure data consistency, completeness, compatibility, shareability and evolvability. Porter and Perry (1984:40) advocate that enterprise-wide data "consistency" can only be achieved when a "top-down" approach to data modelling is followed by organisations as they explain:

*The needs and realities of large enterprises may require them to implement multiple databases to support the operational aspects of IRM. All that IRM demands with respect to such multiple implementations is consistency of the data contents across the enterprise. This consistency may be best accomplished by treating all data as part of a global logical definition of data.*

**FIGURE 4.4 COMPONENTS OF THE CONCEPTUAL DATA MODEL**



Source: McFadden & Hoffer, 1991. Database Management. California: The Benjamin/Cummings Publishing Company, Inc. (page 172).

The whole concept of logical integrity, in the context of this study, revolves around this aspect of ensuring that the physical data base represents in "structure" and "meaning" the "global logical definition of data" (conceptual data design). Structure and meaning, in terms of the logical integrity aspect of data, that need to be maintained, relate to the components that comprise a conceptual data model (i.e., entities, relationships, etc.), as depicted in Figure 4.4.

Auditor evaluation of the controls which ensure that the logical integrity of data is maintained, is the focus of the discussions of this section. The aim of the total audit effort, however, is to predict the reliability of information systems based on the evaluation of the overall control system which ensures data (logical and physical) integrity as well as process (application) reliability. The assumption is made that the reliability of the infrastructure (networks, hardware and system software, etc.) that support the information system falls outside the scope of this evaluation.

#### **4.5.2 The impact of data management on the traditional systems development audit process per life cycle phase**

For the purpose of this study in determining the impact of a data management approach on the systems development audit approach the development life cycle phases, to be discussed, will be based on an information engineering (IE) approach to systems building since an IE methodology incorporates a data base development life cycle as Martin (1986:32) implies:

*The term information engineering refers to the set of interrelated disciplines which are needed to build a computerized enterprise based on data models.*

The view is expressed that an information engineering emphasis to systems development caters for the shortcomings relating to a proper data management approach in traditional systems development life cycle methodologies.



The phases that will be discussed will not include the "information strategy planning" phase, since, it is believed, that this phase is actually a "pre-phase" to systems development. An "initiation" phase will, however, be added as this phase provides transition from systems planning to systems development projects.

The major life cycle phases of systems development in context of a data management approach (to systems building), therefore are:

- initiation
- analysis
- design
- construction

#### 4.5.2.1 Initiation phase

##### PHASE DESCRIPTION

This phase consists of "planning" tasks to be performed at a project level. It is differentiated from the information strategy phase or systems planning, which defines the high-level models, the overall goals and mission of the organisation. The objective of this phase is to provide a means of transition from systems planning to project initiation activities that *provide the foundation for meeting the overall goals* of the organisation, explains the Institute of Internal Auditors (1991b:5-52). They expand:

*The business functions defined in the systems plan divide the enterprise into logical business areas. These areas are further analyzed during project initiation to identify the information systems that are needed to support each area. Other activities consist of defining high-level projects for the business area, determining business needs, performing feasibility study, defining target information systems, selecting an implementation strategy, developing an implementation plan, and establishing a project plan.*

## THE AUDIT APPROACH TO THE INITIATION PHASE

Kothari (1988:60) asserts that this phase does not warrant excessive audit involvement. He believes that the audit emphasis during this phase should mainly be that of planning the audit scope.

## THE IMPACT OF A DATA MANAGEMENT EMPHASIS ON THE INITIATION PHASE

Gilhooley (1990:8) indicates that the interaction of data management during this phase is minimal. The only interaction emanates from data management's involvement during the previous phase in compiling the corporate or overall data model.

## THE IMPACT OF A DATA MANAGEMENT EMPHASIS ON THE AUDIT APPROACH TO THE INITIATION PHASE

It is believed that since a data management approach does not impact the initiation phase that the audit approach will stay the same whether or not a data management emphasis is employed.

### 4.5.2.2 Analysis phase

#### PHASE DESCRIPTION

During this analysis phase the user requirements are defined in terms of business processes, information requirements and the relationships between them.

Martin (1990:197) emphasises that diagrams and matrices are utilised to model and record the data (information requirements) and activities (business processes) of the organisation. The diagrams and matrices should be depicted in such a way that management, end users and data processing professionals can easily understand their content and to increase

communication among these groups.

The high-level descriptions of security, control and audit requirements have to be addressed during this phase. The data, process and supporting descriptions as well as the definition of the technology environment have to be organised in a logical manner of how the system will ultimately be implemented. (IIARF 1991b:5-52).

Martin (1990:201) claims that *fully automated diagramming tools with an encyclopedia* (CASE tools) are to be employed for efficient analysis and design.

With the huge volumes of data associated with the overall data base environment, it is essential to maintain a central index encyclopedia, or dictionary, of all the data items or elements within the environment. This index is often called a "repository" and consists on a conceptual level of all the data items and their relationships to each other, their location in the storage media and which programs or sub-schemas utilise them. The concept of this repository can be extended to include human-readable definition of each term and an indication of the ownership of each item.

This view is expressed that, such a repository is necessary to control the overall environment in terms of new additions, maintenance and other changes of data definitions. The repository could theoretically be maintained manually, but practical considerations dictate the use of computer or automated tools such as software packages as Martin (1990:201) indicated. For the purpose of this study the fact is acknowledge that a data repository should be automated to effectively and efficiently control the meta-data (data about data) should also serve as a "knowledge base" for the organisation as explained in Section 3.2.4.2.

#### THE AUDIT APPROACH TO THE ANALYSIS PHASE

Kothari (1988:59) propagates that the auditor's responsibilities during this

phase should include:

- an evaluation of the adequacy of the application controls over transaction processing and the *environmental* controls surrounding it
- steps to ascertain whether the functional specifications reflect all user requirements accurately
- steps to determine whether project activities are carried out in a controlled, efficient and competent manner

It is also important that the auditor begin thinking about the future audit requirements to be built into the system and audit approach to auditing within the new system environment during this phase (Wysong 1983:37).

#### THE IMPACT OF A DATA MANAGEMENT EMPHASIS ON THE ANALYSIS PHASE

Martin (1990:245) advocates that a vital part of the analysis phase, in terms of an information engineering methodology, should be allocated to the creation of a model of the data as it forms the foundation on which applications will be built.

During this phase a model is compiled that reflects the business which should not include any data resulting from systems design. It is an iterative process that leads to increasing amounts of detail. The process includes the definition and documentation what the business (users) require in the data base to satisfy present and future information needs (McFadden & Hoffer 1991:171). The data model at this stage normally contain many types of data (entities). The activities that lead to the compilation of the conceptual model during this phase include (IIARF 1991b:5-23):

- review of deliverables and work products for systems planning activities

- identification of project or subject area(s) pertinent to the system(s) being developed
- identification, description, and documentation of pertinent business entities (persons, places, things, or concepts of lasting interest to the business)
- identification of relationships and rules that govern the association between business entities (e.g. purchase order may be issued to one and only one vendor; each purchase order will have at least one order line item, but it may have a maximum of forty-five line items; no purchase order may be placed that orders parts for more than one customer order)
- discovery of candidate keys or those attributes of an entity that uniquely identify occurrences of that entity (a key is an attribute or combination of attributes that can be used as an identifier of entities, because no pair of entities share the same value of that key)

During this phase the data administrator is mainly responsible for ensuring that:

- the data model is reconcilable with the data planning (architecture) models compiled during the information strategy planning phase (Gilhooley 1990:6)
- the data definitions and models are properly documented (Ravindra 1987:2; Sibley 1983:2; Martin 1986:50)
- there is agreement among the users about the definitions and format of the data depicted in the data models (Martin 1986:47)
- conflicts about incompatible representations of data are resolved (BCS 1989:38)

## THE IMPACT OF A DATA MANAGEMENT EMPHASIS ON THE AUDIT APPROACH TO THE ANALYSIS PHASE

By participating in this phase, the auditor should be able to determine whether the conceptual models include an *adequate network of internal controls* and sufficient and identifiable audit/management trails. The auditor should also be in the position to determine that the user's requirements have appropriately been addressed and the output products of this phase are adequate to meet those requirements (Wysong 1983:36).

The output of this phase in the context of a "data management" environment takes the form of conceptual process and data models with properly defined entities, relationships and business rules which provide a concise and clear means to the auditor for obtaining an understanding of the proposed system. The view is therefore expressed that the auditor's task, of evaluating the adequacy of controls at this stage of development and determining whether user requirements are addressed, is considerably eased as proper documentation, diagrams and models of the data (as well as the process component) are compiled and maintained during this phase.

It is believed that, in determining whether the user's information requirements have appropriately been addressed, the auditor will have to take a more "holistic" approach as the conceptual data design as well as the data definition at this level will have to be audited in the context of the higher-level enterprise model including its information/entity descriptions and definitions.

### 4.5.2.3 Design phase

#### PHASE DESCRIPTION

This phase of system development involves adapting the conceptual design into an effective system. The design phase translates the conceptual models into a blueprint of the processing functions that the system will perform to

meet user needs (IIARF 1991b:5-52).

The design phase is the most important phase of systems development. During this phase the details pertaining to systems access, business and processing control procedures and auditability requirements are incorporated into the design. Data and process diagrams and specifications are organised into a system that fits within the constraints of the available platform technology (hardware, software and network). (Nasuti 1987:7; IIARF 1991b:5-52)

#### THE AUDIT APPROACH TO THE DESIGN PHASE

Kothari (1988:59) is of the opinion that the main task of the auditor, during this phase, is that of evaluating the detailed system design for internal controls and consistency with the functional specifications.

Nasuti (1987:7), in support of Kothari, emphasises that from an operational audit perspective, the auditor must also ensure that the system is designed as intended by the user and management. In doing so, he believes the auditor should review the detailed system design specifications to determine whether they:

- reflect approved user requirements
- meet applicable corporate policies, standards and practices
- provide adequate internal controls, security, back-up and recovery, requested audit features and audit trail

#### THE IMPACT OF A DATA MANAGEMENT EMPHASIS ON THE DESIGN PHASE

During the design phase the data model created during the analysis phase is developed into more detail and fully normalised. As data is applied to application systems further data analysis may uncover new entities and

relationships. The results of the analysis should be fed into the higher-level models to validate model integrity (Meador 1989:5).

The activities that lead to the logical data model include (IIARF 1991b:5-25):

- convert business entities to data entities/data groups through the logical groupings of data that will eventually be stored and managed in some form by a computer
- define new data entities (e.g., characteristic, supertype, and subtype entities)
- define data elements associated with data entities and data element attributes
- define primary and foreign keys; a primary key is a particular candidate key selected to distinguish occurrences of a data entity; a foreign key is a data element or combination of data elements that references the primary key of a related entity; a foreign key must carry the same value as its corresponding primary key or have a missing (null) value; relationships established by foreign keys are used to define referential integrity constraints
- specify availability and concurrency requirements; the requirements include the recovery and restart times and techniques that usually are derived from the needs of the processes using the data
- define user data views; the user data views are defined to provide the subsets of data for each process; this helps provide better integrity, security, and performance
- define data relationships or the associations or connections present among data entities; relationships are determined by the business rules and relationships of the subject area; three generic kinds of



relationships that can exist among occurrences of entities include the following:

- one entity can relate to zero or one occurrence of another entity (e.g., one employee may be assigned one office)
  - one entity can relate to zero, one, or many occurrences of another entity (e.g., one purchase order may have many orderline items)
  - one entity must relate to one but may relate to many occurrences of another entity (e.g., a purchase order must designate at least one contract)
- define data volumes and determine update (add, change, or delete) frequencies for each data entity
  - normalise and refine data entities into third normal form; normalisation is a data design technique used to minimise potential update (add, change, or delete) problems that can occur when a data element is incorrectly placed in an entity (data group) or when the same data element is found in multiple entities; the designer's aim, through the use of the normalisation technique, is to have one fact in one place; each data element of an entity is to be dependent on the entire key and nothing but the key; successful use of the normalisation technique requires a thorough understanding of the business rules and relationships that govern the subject area; the update anomalies that can occur are not due to deficiencies in database management system functions or operability; normalisation results in a set of restructured or rearranged data entities

Meador (1989:6) explains that the purpose of the fully normalised model (canonical schema) is to:

*... define the information requirements in terms of a logical data*

*base. No data item is duplicated in the model, it is in full relational form. Such a format facilitates the accessibility of all data to all business processes and promises the most flexibility in the use of data. It is through this model that systems data sharing and application integration can occur.*

Gilhooley (1990:8) points out that the systems development manages the development activities during this phase. However, data management activities assigned include:

- the responsibility to ensure the model and the data definitions are reconcilable with the conceptual model and overall data architecture as well as the data definitions compiled during the analysis and data architecture compilation phases
- further refinements to the data element definitions and data structures (models)
- modifications to the data dictionary and corporate data model
- approval of the detailed data design

The data administrator therefore ensures that all the data during this phase are well defined, analysed and classified into a third normal form view. Gilhooley (1990:8) believes that is important that the data base administrator also approves the data design at this stage.

#### THE IMPACT OF A DATA MANAGEMENT EMPHASIS ON THE AUDIT APPROACH TO THE DESIGN PHASE

The view is expressed that a data management emphasis to the design phase extends the "holistic" audit approach taken during the analysis phase further since the logical design of the data base will have to be audited in the context of the previous phases conceptual design as well as the higher-level enterprise model.

The formalisation of the data requirements in terms of a logical data base design, in a shared data base environment also requires a well balanced enterprise-wide perspective to be taken regarding the data structure and definition as Davis and Weber (1983:77) explains:

*The design of the database must satisfy multiple users. Moreover, a design error affects multiple users; thus, compared to single file systems, the expected losses from an error or irregularity increase. Somehow a global perspective of the design process must be adopted.*

The quality of the data base will depend upon rigid enforcement of consistent data definitions. At its most basic this means that everyone must refer to a data item by the same name, otherwise duplication of data within the data base will probably result which may lead to serious inaccuracies apart from loss of efficiency and a generally weakened opportunity to integrate applications. In practice this could be more of a problem than might appear because what could be referred to as "account balance" by one department could be called client balance, customer balance or balance in the other departments. Controls must be in place to ensure that like items are held on the data base once only and that no items are deemed to be the same unless they really are. The auditor should confirm in his audit that data definitions and design are used by the different users in accordance with one common definition, otherwise misleading or erroneous management information may be produced. (Chambers & Court 1991:181)

It therefore follows that, in a data base environment where the same data element can be available for many purposes, it requires a broader view of the data in terms of its meaning, structure and use and correspondingly a more holistic audit approach. As auditors have always been concerned with clear documentation of the source and use of data they now will have to expand their perspective to consider the multiple relationships that may exist.

Nasuti (1987:7) stresses that, from an operational viewpoint, the auditor's involvement in the design phase is crucial. The opinion is therefore

expressed that the auditor's involvement during this phase, as it relates to the data aspects of the design in ensuring that the appropriate controls are in place, is vital. The auditor must, however, take a very broad view on the logical integrity aspects of the data as it will be important to ensure that the data definitions and structures serve not only the specific application scope but the overall data base scope.

#### 4.5.2.4 Construction phase

##### PHASE DESCRIPTION

During this phase the system is built, and the activities performed normally include (IIARF 1991b:5-52):

- program design
- program coding
- program testing
- data base construction
- data base population
- systems testing

Testing fulfils a vital role during this phase. The Institute of Internal Auditors explains (IIARF 1991b:5-52, 5-53) the following testing needs to be performed:

*Initial testing at both the unit and system level, is performed in the development environment. User acceptance testing is typically performed in a controlled test environment and may be managed by the IS quality assurance staff. Changes in platform-specific facilities should be unit and system tested by appropriate personnel to ensure the stability and reliability of the data center environment.*

##### THE AUDIT APPROACH TO THE CONSTRUCTION PHASE

The audit effort during this phase will focus on the two main activities

performed – coding and testing.

Kothari (1988:59) believes that the main task of the auditor, in terms of the coding aspect of the phase, would be that of evaluating the change controls over code delivery. The audit tasks to be performed (relating to the testing phase) should include:

- a review of the systems testing process, including adherence to the test plan
- a review of how the problems discovered during testing were resolved
- re-performance testing, if required

#### THE IMPACT OF A DATA MANAGEMENT EMPHASIS ON THE CONSTRUCTION PHASE

During this phase the logical data model, completed during the previous stage, is split up into two distinct models: a model in terms of the *activities that relate to data storage* and a second model that relates to *data manipulation and use*. Constructing these physical data base models is in some respects the most difficult phase of developing the overall data architecture, comments Meador (1989:7).

The activities performed during this phase that lead to a physical data model include (IIARF 1991b:5-26, 5-27):

- specification of the appropriate physical file organisation and/or physical data base storage structure and the translation of data entities into DBMS specific storage structures for implementation (e.g., segments, tables, etc.)
- definition of physical data storage characteristics (e.g., block size, page size, area size, physical record type, overflow storage area,

definition of indexes, pointers, links, and compression techniques)

- refinement of file and/or data base storage volume estimates, update frequencies, and capacity requirements
- evaluation of physical data structure efficiency via transaction path analysis
- distribution of the data across multiple platforms
- refinement of physical data structure for greater performance and system throughput, denormalisation (reverse normalisation) for improved retrieval efficiency
- definition of data base schema and subschema

The data base administrator is responsible for the actual physical data base design that provides the foundation for the integration of the data resources residing in the computer (Ravindra 1987:6; Durell 1984:7).

The implementation of physical data base structure during the construction phase results in extra testing during the testing sub-phase as the newly developed application needs to be tested against a test data base, along with other application programs that are in production mode (Plagman 1982:8, Gilhooley 1990:4).

#### THE IMPACT OF A DATA MANAGEMENT APPROACH TO THE AUDIT APPROACH TO THE CONSTRUCTION PHASE

The view is expressed that the auditor's approach to the construction phase will not be severely impacted by a data management emphasis. The auditor will, however, be required to ensure that the newly developed information system is adequately tested against a test data base along with other programs that are in production mode and in addition that the data base

administrator as well as all the other concerned parties have "signed off" the system before the introduction of the system into production.

#### **4.6 SUMMARY**

It was established that the objectives of the systems development life cycle audit process and systems auditing in general, do not change when a formal data management approach to information systems management is introduced into the organisation.

It was, however, concluded that the audit of information systems in a data base environment requires a more holistic audit approach to the evaluation of the control system. This holistic approach requires the auditor to expand the scope of his information system audit to incorporate an evaluation of the controls that ensure not only the integrity of the particular information system's data but also that of the entire data base.

The research conducted illustrated that it is advisable to conduct the audit of the application systems controls during the various stages of development and for internal and external auditors to cooperate in conducting these audits.

The research illustrated that there is a definite trend toward the migration of controls from the application level to the general systems environment in the data base arena.

Following on this identified trend of migration of controls to a "higher level" and the requirement for a more holistic audit approach, it is concluded that auditors need to take a new look and place more emphasis on the controls and especially the organisational controls which ensure the reliability of the overall data base of an organisation. In this regard this research study has illustrated that a formal data management function is a keystone organisational control to ensure the reliability of an organisation's information.

## CHAPTER 5

### CONCLUSIONS AND RECOMMENDATIONS

#### 5.1 INTRODUCTION

*Many observers of trends in business believe that the organisations that will excel in the 1990s will be those that manage information as a major resource and structure it as efficiently as they do other resources. Of, course, information can be an asset only if it is accurate and available when needed, which can occur only if an organization purposely organizes and manages its data. The database has become the standard technique for structuring and managing data in most organizations today. (McFadden and Hoffer 1991:3)*

The management of an organisation's information resources has become a key process within most organisations today and the areas of data administration, data architecture compilation, and data base-oriented systems all present unique management issues, as well as audit and control considerations. (IIARF 1991b:5-50)

As organisations are increasingly starting to introduce data base technology and data management practices into business, it is becoming more and more important for auditors involved in information systems audits to understand how the integrity of data and information can be impacted as the result of the new technology and data management practices.

The professional auditing bodies require auditors to keep abreast with the technological developments in the field of information management to enable them to develop skills and appropriate audit strategies as well as audit procedures in meeting the demands placed on auditing by the ever evolving technologies.

The South African Institute of Chartered Accountants, in their Code of Professional Conduct: ET 040, titled Professional Skill, updated during 1992,



stipulates the requirement that an auditor has to *maintain his professional knowledge and skill at a level sufficient to ensure his client or employer receives the advantage of competent professional service, based on up-to-date developments in practice, legislation and techniques* (SAICA 1992:para .02).

In line with ET 040, the Standards for the Professional Practice of Internal Auditing, Standard 220: Knowledge, Skills, and Disciplines, issued by the Institute of Internal Auditors during 1978, stipulates that (IIA 1978:12):

*The internal auditing department should possess or should obtain the knowledge, skills, and disciplines needed to carry out its audit responsibilities.*

The purpose of this research study is outlined in Chapter 1 as that of providing auditors of all descriptions, both internal and external, with an understanding and knowledge of the impact that modern data base technology and data management practices have on the process of information systems and systems development auditing. It is believed that without a proper understanding, auditors can not perform an effective and a meaningful audit in the corporate data base environment.

Two sub-problems were identified in Section 1.5 dealing with the effect of corporate data base technology and data management practices on the traditional systems development process and systems auditing respectively. In order to address these sub-problems the following two objectives were set:

- firstly, to ascertain the effect that corporate data base technology and data management practices have on the traditional systems development process
- secondly, to determine the impact of data base technology and

commensurate data management practices on the systems development life cycle audit process and systems auditing in general

In meeting these two objectives a literature study of authoritative literature on information technology, information systems, computer science and auditing was undertaken.

In this chapter it is proposed to present:

- a summary of the research findings pertaining to the two objectives
- the conclusions derived from research findings
- the recommendations for additional research required

## **5.2 RESEARCH FINDINGS PER IDENTIFIED OBJECTIVE**

### **5.2.1 Objective 1:     To ascertain the effect of corporate data base technology and data management practices on the traditional systems development process**

In establishing the above effect the most significant aspects and components of a data base environment as well as the concept data resource management were researched within the context of an information resource management (IRM) approach to systems development as this latter approach is the guiding principle behind the development of data bases and data resource management. The purpose was to obtain an overall understanding of these "intertwined" concepts as a prerequisite for ascertaining the effect of the set concepts on the traditional systems development process.

#### 5.2.1.1 Obtaining an understanding of data base technology and data management practices within an overall IRM approach to information management

It was established that IRM is the concept that information is a major corporate resource that must be managed using the same basic principles that are used to manage the organisation's other assets, such as human resources and finances.

The research findings indicated that a proper resource management approach to information should be focused on the organisation's information requirements in support of its goals, critical success factors and plans which result in a new approach to information system design – a shift from an operational systems emphasis toward systems that focus on the critical tasks and decisions made by managers.

The research illustrated that the challenge of an IRM emphasis to information systems is to deliver the right information to the right people at the right time which requires:

- a proactive approach to information management
- strategic information planning and the development of mission critical systems incorporating an information-engineering type of approach
- an information technology architecture
- centralised information management

Most authoritative observers emphasise that IRM includes the management of the underlying data resource which should also be treated and managed as a critical corporate resource. The view is expressed that data, in this sense, also requires to be managed pro-actively and to be strategically planned in terms of a properly defined data architecture that forms part of

an overall information technology architecture.

Meador (1988:1) postulates that the main purpose of data base technology is to provide the organisation with a set of tools that work together to manage the organisation's data resources.

It is against this background of research findings and perspective obtained, that the impact of the set technology and management practices was researched to determine their impact on traditional system development.

#### 5.2.1.2 Determining the effect of a data management approach on the traditional systems development process in the context of a data base environment

The research conducted illustrated that the influence of data management on the traditional systems development process is quite profound in the sense that a data management approach results in:

- a data-driven emphasis to systems development
- a (separate) data base development life cycle

It was established that a data-driven approach require an organisation-wide conceptual data representation or model from which the logical and physical design of data base(s) of the organisational units can be derived – an approach that necessitates a separate data base development life cycle that has to proceed in parallel with systems design. The data base development life cycle encompasses the activities that range from the identification of end-user requirements to the final physical arrangement of the data values.

The following summary which lists the main differences between a data-driven (data base development) life cycle and a traditional systems development life cycle clearly demonstrates that the first objective has been met:

- the data base development life cycle focuses on data specific deliverables throughout all the life cycle phases, whereas the traditional systems development life cycle only delivers system specific requirements and logical data views during the first two phases
- a particular system's data model is (always) addressed within the scope of the overall corporate data model in terms of the data base development life cycle approach to ensure data consistency across applications in contrast to the particular logical data views which are derived within the context of the system requirements for the traditional systems approach
- data becomes the "foundation block" for developing systems within the context of a data base development approach compared to the traditional development approach where systems are developed mainly to respond to processing requirements

**5.2.2 Objective 2:     To determine the impact of data base technology and commensurate data management practices on the systems development life cycle audit process and systems auditing in general**

*Internal auditors and accountants rely on internal control to ensure the integrity of a system and the data it maintains. Internal control provides management with reasonable, but not absolute, assurance that financial and other resources are safeguarded from unauthorized use or disposition; transactions are executed in accordance with authorizations; financial and statistical records and reports are reliable; applicable laws, regulations, and policies are adhered to; and resources are efficiently and effectively managed.*  
(Brown 1991:1)

Mair, Wood and Davis (1986:8) explain that the internal controls of information systems need to be evaluated to predict their reliability, an approach that is commonly known as compliance auditing.

Many observers (including some of the professional auditing bodies) propagate that this audit process should commence before a system is up and

running and auditors should therefore conduct their audit while the system is being designed and developed especially when large, complex and highly integrated systems are encountered.

Davis and Perry (1982:21) believe that the auditing effort in terms of the evaluation of systems controls should be considered in its "totality" and not split between internal and external auditing. The research findings indicated that many observers, in supporting their view, are of the opinion that internal auditors should carry out the compliance audit tests and report on findings to the external auditor. In this way duplication of effort is avoided which results in improved coordination of the total effort.

Ford (1987:8) further comments that:

*External auditors are going to have to accept that with increasing system complexity, more reliance is going to have to be made of the work of internal auditors ...*

Within the scope of this study the research focused only on the particular logical reliability or integrity aspect of the data component of information systems which is commonly known as "logical data integrity". The Institute of Internal Auditors (IIARF 1991b:5-13, 5-14) emphasises that one must, however, remember that the overall systems reliability depends on the integrity (or reliability) of the process which transforms "raw" data into information, as well as the integrity of the raw data itself.

Logical data integrity in terms of the scope of this study is defined as the measures taken to ensure that the physical data base structure and meaning of the data model used to design it.

The research conducted indicated that an integrated data base environment with the specific objective of shared data, typified by a number of users making use of the common data base, necessitates the establishment of data integrity as an important control objective of the organisation and its

management.

In determining the impact of data management on the auditing process as it relates to the evaluation of appropriate controls to ensure logical data integrity, in the context of the scope of this study, it was established that an effective data management function fulfils a key role to ensure the integrity of an organisation's data is maintained and as such becomes an essential control on which the auditor can rely in terms of his evaluation of system of internal controls.

It is concluded that the objective of determining the impact of a data management approach on the systems audit and systems development audit process has been met.

### **5.3 CONCLUSIONS**

It was established that the auditing objective of determining the reliability of systems is not affected in any way by the introduction of data base systems and data management practices. The introduction of data base technology, however, tends to increase the exposure that may result from poor control since the data base itself contains data from many sources which is also shared by many users.

A formal data management function fulfils a key role in this regard, in terms of ensuring that the integrity of the overall corporate data base stays intact.

In considering the research findings, the hypothesis has been proved in that the study results clearly illustrated that the use of corporate data base technology and data management practices introduces a new consideration to the systems development process in that the auditor must adopt a more holistic approach to auditing of information systems.

It was also illustrated in the findings that this holistic audit approach requires the auditor to expand the scope of his information systems audit to

incorporate an evaluation of the controls that ensure not only the integrity of the particular information systems data, but also that of the entire data base of the organisation.

This particular finding illustrates that the emphasis outlined in the hypothesis is correct in that the audit of any specific information system should be viewed in the context of the bigger data base system and as such should be viewed as one of many sub-systems that utilises and updates the corporate data base in support of the mission of the organisation.

#### **5.4 RECOMMENDATIONS FOR FUTURE RESEARCH**

The research conducted as it relates to the role of the data management function in ensuring the logical integrity of data did not specifically address the impact of computer aided software engineering (CASE) technology.

Although reference in the study was made to CASE being a component of the data base environment, its capabilities supporting data management in terms of creating a knowledge base of conceptual data designs and automated facilities to generate physical data base (physical structures) require further research.



## BIBLIOGRAPHY

### BOOKS

Anthony, R.N. 1965. Planning and control systems : A framework for analysis. In Division of Research, Graduate School of Business Administration. Boston: Harvard University Press.

Bainey, K.R. 1986. A Practical Approach to building an information systems plan. In Auerbach systems development management. 33-03-20. New York:Auerbach : 1-17.

Biggs, C.L., Birks, E.G. & Atkins, W. 1980. Managing the systems development process. New Jersey:Prentice-Hall.

British Computer Society: Data Management Specialist Group. 1989. Information management and planning. Gower Technical.

Brown, N. 1991. Implementing integrity in automated applications. In Auerbach EDP auditing. 75-01-25. New York:Auerbach: 1-11.

Butler Cox Foundation. 1988. Managing the evolution of corporate databases. Butler Cox Foundation : Research report 64, June 1988. London:Butler Cox.

Chambers, A.D. & Court, J.M. 1991. Computer auditing. London:Pitman Publishing.

Curtice, R.M. 1990. Guidelines for understanding and developing data-driven systems. In Auerbach data base management. 22-01-30. New York:Auerbach: 1-13.

Date, C. J. 1977. An introduction to database systems. 2nd Edition. Massachusetts: Addison - Wesley.

Davis, G. 1968. Auditing and EDP. New York:American Institute of Certified Public Auditors.

Davis, K.W. & Perry, W.E. 1982. Auditing computer applications: A basic systematic approach. New York:John Wiley & Sons.

Davis, G.B. & Weber, R. 1983. Auditing advanced EDP systems: A survey of practice and development of a theory. The management Information Systems Research Center, University of Minnesota.

Dawley, D.L. 1983. What auditors should know about data processing. Ann Arbor, Michigan: UMI Research Press.

Duffy, N.M. & Assad, M.G. 1980. Information management: An executive approach. Cape Town:Oxford University Press.

Durell, W.R. 1985. Improving data consistency, reliability, and accuracy. In Auerbach data base management. 22-05-08. New York:Auerbach : 1-12.

Earl, M.J. 1989. Management strategies for information technology. New York:Prentice-Hall.

Fleming, C.C. & Von Halle, B. 1990. Logical data modeling. In Auerbach data base management. 22-30-21. New York:Auerbach: 1-14.

Gallegos, F. & Dow, D. P. 1987. Key review points for SDLC audits. In Auerbach EDP auditing. 76-01-25. New York:Auerbach : 1-15.

Gilhooley, I.A. 1985. A Productive Approach to Systems Development. In Auerbach systems development management. 31-01-01. New York:Auerbach: 1-8.

Gilhooley, I.A. 1986. The EDP Internal/external auditor relationship. In Auerbach EDP auditing. 72-01-60. New York:Auerbach: 1-12.

Gilhooley, I.A. 1990. Managing the corporate information resource. In Auerbach data base management. 24-02-10. New York:Auerbach: 1-12.

Griffin, R.W. 1987. Management. Boston:Houghton Mifflin.

Holloway, S. 1988. Data administration. Vermont:Gower.

Knight, A.V. & Silk, D.J. 1990. Managing information: Information systems for today's general management. London:McGraw-Hill.

Mair, W.C., Wood, D.R. & Davis, K.W. 1986. Computer control and audit. Altamonte Springs, Florida:The Institute of Internal Auditors.

Martin, J. 1976. Principles of data base management. Englewood Cliffs, New Jersey: Prentice-Hall.

Martin, J. 1980. Managing the data base environment. Lancashire:Savant Institute.

Martin, J. 1981. An end-user's guide to data base. New Jersey:Englewood Cliffs, Prentice-Hall.

Martin, J. 1986. Information engineering. Lancashire:Savant Research Studies.

Martin, J. 1989. Information engineering: Introduction (Book I). New Jersey:Englewood Cliffs, Prentice-Hall.

Martin, J. 1990. Information engineering: Planning and analysis (Book II). New Jersey:Englewood Cliffs, Prentice-Hall.

McFadden, F.R. & Hoffer, J.A. 1991. Database management. California:The Benjamin/Cummings Publishing Company, Inc.

Meador, J.G. 1988. Planning the data base environment. In Auerbach data base management. 23-30-15. New York:Auerbach : 1-13.

Meador, J.G. 1989. A data architecture framework : The business perspective. In Auerbach data base management. 21-01-01. New York:Auerbach: 1-13.

Meltzer, M.F. 1981. Information : The ultimate management resource - how to find, use, and manage it. New York:AMACOM.

Moeller, R.R. 1988. Long-range information systems planning. In Auerbach Systems Development Management. 33-03-30. New York:Auerbach:1-16.

Morton, M.S.S. 1991. The corporation of the 1990s: Information Technology and organisational transformation. New York:Oxford University Press.

Moss, C.L. 1983. The information systems plan and the data base plan. In Auerbach data base management. 21-02-04. New York:Auerbach: 1-11.

Moss, C. L. & Duè, R. T. 1988. Security and control opportunities in data base design. In Auerbach data base management. 23-30-30. New York: Auerbach : 1-12.

Naisbitt, J. 1984. Megatrends : Ten new directions transforming our lives. London:Macdonald.

Nasuti, F.W. 1987. Auditing during the development life cycle. In Auerbach EDP auditing. 76-01-15. New York:Auerbach: 1-12.

Oxford Dictionary 1987. The South African pocket Oxford dictionary. Cape Town:Oxford University Press.

Perry, W.E. 1985. Systems auditability and control reports. In Auerbach EDP auditing. 71-03-02. New York:Auerbach: 1-8.

Perry, W.E. 1986. Auditing applications that use a data base. In Auerbach EDP auditing. 74-02-40. New York:Auerbach: 1-16.

Perry, W.E. 1987. Ensuring the integrity of the data base. In Auerbach EDP auditing. 75-02-40. New York: Auerbach : 1-16.

Peters, T.J. & Waterman, R. H. 1982. In search of excellence - Lessons from America's best-run companies. New York:Harper & Row.

Plagman, B.K. 1982. Systems development in a data base environment. In Auerbach data base management. 23-02-01. New York:Auerbach: 1-12.

Porter, W.(Jr) & Burton, J.C. 1971. Auditing: A conceptual approach. California:Wadsworth.

Porter, W. T. & Perry, W. E. 1984. EDP:Controls and auditing. Boston: Kent.

Ravindra, P.S. 1987. Data administration : An old function adapts to its new role. In Auerbach Data Base Management. 22-20-10. New York:Auerbach : 1-13.

Ravindra, P.S. 1990. Managing data from a corporate perspective : The role of the data administration function. In Auerbach data base management. 24-03-30. New York:Auerbach: 1-8.

Roberts, M. B., 1985. EDP Controls. New York: Wiley.

Ross, S.J. 1982. Data administration : Control, structure, and organization. 22-05-04. In Auerbach data base management. New York:Auerbach: 1-11.

Sardinas, J. & Burch, J.G. 1981. EDP Auditing: A primer. New York:John Wiley & Sons.

Sibley, E. H. 1983. An Introduction to data base management – Part II. In Auerbach data base management. 21-01-02. New York: Auerbach : 1-13.

Sibley, E.H. 1983a. An introduction to data base management – Part I. In Auerbach data base management. 21-01-02. New York:Auerbach: 1-15.

Synnott, W.R. & Gruber, W.H. 1981. Information resource management: Opportunities and strategies for the 1980s. New York:A Wiley-Interscience Publication.

Taylor, D.G. 1983. Recommendations for a statement of generally accepted auditing practice on using the work of an internal audit function. The Department of Accounting:University of Cape Town.

Thompson, J.P. 1989. Data with semantics: Data models and data management. New York:Van Nostrand Reinhold.

Watne, D.A. & Turney, P.B.B. 1990. Auditing EDP systems: Second edition. New Jersey:Prentice-Hall.

Weber, R. 1982. EDP auditing : Conceptual foundations and practice. New York : McGraw-Hill.

Weldon, J. 1990. Trade-offs in data base design. In Auerbach data base management. 22-02-10. New York: Auerbach : 1-11.

Wiseman, C. 1988. Strategic information systems. Illinois:Irwin, Homewood.

Wysong, E.M. 1983. Information systems auditing. Amsterdam:North-Holland.

## JOURNALS

Appleton, D.S. 1983. Law of the data jungle. Datamation, October:225-230.

Appleton, D.S. 1985. The technology of data integration. Datamation., November 1:106-116.

Appleton, D.S. 1986. Rule-based data resource management. Datamation, May 1:86-99.

Borthick, A.F. 1986. Audit implications of information systems. The CPA Journal. April:40-46.

Brancheau, J.C. & Wetherbe, J.C. 1987. Key issues in information systems management. MIS Quarterly, March: 22-29.

Bryce, M. 1987. The IRM Idea. Datamation, April 15: 89-92.

Chen, Y. 1983. A new era in external / internal auditor cooperation. CA Magazine (Canada), September:74-76.

Curtice, R. M. 1986. Getting the database right. Datamation, October 1: 99-104.

Durell, B. 1984. Keepers of the corporate data. Computerworld, July 30:1D5-1D12.

Etheridge, J. 1986. Banking on a safety net. November 1: 64.21-64.22. Datamation.

Ford, J. 1987. How do we train and who shall be responsible? Audit Update, February/March:8.

Gilhooley, I.A. 1984. Auditing system development mythology. Edpacs, XII (1):1-8.

Jenkins, B.G. 1981. The future of computer auditing. The South African Chartered Accountant. May:211-216.

Kothari, N. 1988. Auditing's role in systems development. CA Magazine (Canada). October: 55-60.

Mach, R. 1990. Information engineers wanted to built enterprise models. Software Magazine, September:38-51.

Mendus, S.L. 1986. Computer systems development : Why not let auditors help? Journal of systems management., May vol. 37(5):36-40.

Nolan, R.L. 1979. Managing the crisis in data processing. Harvard business review, March/April:115-126.

Percy, T. 1986. My data, right or wrong. Datamation, June 1:123-128.

Percy, T.P. 1986. My data, right or wrong. Datamation. June 1:123-128.

Reilly, R. F. 1986. The effect of SAS No. 48 on financial and EDP auditors. Edpacs, June: 6-9.

Rollier, B. 1982. Two teams, one goal. Datamation. April. 183-184.

Schatz, E. 1988. A high-technology twist on the Reagan revolution. Datamation, June 1:41-52.

Stoneham, N.W. 1979. Data base auditing - a new complexity : A different means to the same end. CA Magazine:May: 32-38.



Von Halle, B. 1992. Database Design. Database Programming & Design, February:11-13.

Wingate, A.J.K. 1981. New auditing guideline - Auditing in an EDP environment. CA Magazine. October: 76&78.

Zani, W.N. 1970. Blue print for MIS. Harvard business review, November/December: 95-100.

## **AUTHORITATIVE PRONOUNCEMENTS**

AICPA. 1972. Statement on auditing standards No. 1: Codification of auditing standards and procedures. New York: American Institute of Certified Public Accountants (Auditing Standards Executive Committee).

AICPA. 1974. Statement on auditing standards No. 3: The effects of EDP on the auditor's study and evaluation of internal control. New York: American Institute of Certified Public Accountants (Auditing Standards Executive Committee).

AICPA. 1983. Report of the Joint Data Base Task Force. New York: American Institute of Certified Public Accountants.

AICPA. 1984. Statement on auditing standards No. 48: The effects of computer processing on the examination of financial statements. New York: American Institute of Certified Public Accountants (Auditing Standards Executive Committee).

CICA. 1975. Computer audit guidelines. Toronto, Canada: The Canadian Institute of Chartered Accountants.

CICA. 1981. Auditing in an electronic data processing environment. Canadian Institute of Chartered Accountants (Auditing Standards Steering Committee).

IFA. 1985. International auditing guideline No. 20: The effects of an EDP environment on the study and evaluation of the accounting system and related internal controls. International Federation of Accountants (International Auditing Practices Committee).

IIA. 1977. Systems auditability and control study: Data processing auditing practices report. Altamonte Springs, Florida: Institute of Internal Auditors (Stanford Research Institute).

IIA. 1978. Standards for the professional practice of internal auditing. Altamonte Springs, Florida:Institute of Internal Auditors.

IIARF. 1991. Systems auditability and control report:Modules 1-12. Altamonte Springs, Florida:Institute of Internal Auditors Research Foundation.

IIARF. 1991a. Systems auditability and control report – Module 2: Audit and control environment. Altamonte Springs, Florida:Institute of Internal Auditors Research Foundation.

IIARF. 1991b. Systems auditability and control report – Module 5: Managing information and developing systems. Altamonte Springs, Florida:Institute of Internal Auditors Research Foundation.

IIARF. 1991c. Systems auditability and control report – Module 1: Executive summary. Altamonte Springs, Florida:Institute of Internal Auditors Research Foundation.

SAICA. 1986. Accounting systems and Internal control: AU230. In Member's Handbook. Johannesburg: The South African Institute of Chartered Accountants.

SAICA. 1986a. Generally accepted auditing standards: AU010. In Member's Handbook. Johannesburg: The South African Institute of Chartered Accountants.

SAICA. 1989. Auditing and accounting guide: Auditing in a computer environment. Johannesburg: The South African Institute of Chartered Accountants.

SAICA. 1992. Professional skill: ET 040. In Member's Handbook. Johannesburg: The South African Institute of Chartered Accountants.

**DOCTORAL THESISES**

Fink, D. 1991. Evaluation of advanced EDP internal controls: A study of auditor's expertise, audit concerns and EDP skill requirements.  
Pretoria:University of South Africa.

Techavichit, J. V. 1979. Evaluation of internal control in a database environment. Ann Arbor, Michigan, USA: University Microfilms International.

**APPENDIX 1****DATA MODELLING CASE STUDY**

In this Appendix, a case study named Ron's Real Estate Business illustrates the major concepts of data modelling (modeling). The main aspects of data modeling are covered under the following headings:

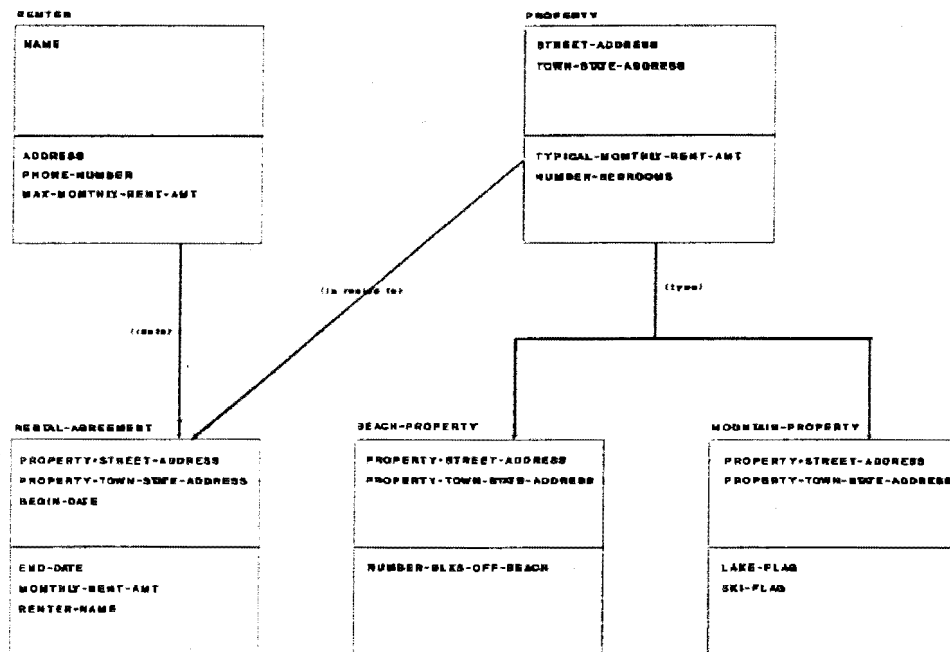
- Ron's Real Estate Business Case Study
- The role of the three-schema architecture
- Steps within logical data modeling
- Critical success factors in logical data modeling
- Benefits and applicability of logical data models
- Summary

**1. RON'S REAL ESTATE BUSINESS CASE STUDY**

Exhibit 1 shows a very simple logical data model of Ron's Real Estate Business. Even knowing nothing about logical data modeling, by looking at Exhibit 1 it is obvious that:

- Ron deals with renters, properties (beach and mountain), and rental agreements – These seem to be the most important objects or concepts (or entities) in the exhibit
- Renters rent rental agreements, and properties are rented to rental agreements – Moreover, properties are of type beach or mountain. The arrows seem to reflect associations (relationships) between the boxes (entities)
- Ron is interested in selected details about each entity – For example, he tracks the name, address, telephone number, and maximum monthly rental amount of each renter. He is concerned with how many blocks away the beach is from a beach property, but not from a mountain

# EXHIBIT 1 A LOGICAL DATA MODEL FOR RON'S REAL ESTATE BUSINESS



- Some attribute names appear above horizontal lines in the boxes (e.g., NAME in the RENTER box). These attributes appear to be the most important or most necessary details about the entity. In fact, these attributes must be present; for example, renters must have names but not always addresses, telephone numbers, or monthly rental amounts
- Some attributes have names containing asterisks (e.g., RENTER\*NAME, PROPERTY\*STREET-ADDRESS, PROPERTY\*TOWN-STATE-ADDRESS). These attributes seem to refer to other entities (or relate them to each other)
- The arrows are drawn with one or two arrowheads at one end and have no arrowheads at the other end—arrowheads presumably carry some meaning

Without any knowledge of logical data modeling, a lot of information can be obtained about Ron's business. Thus, it is obvious that two benefits of the

logical data modeling technique are that it is simple to understand, because it uses uncomplicated diagrams and it expresses many facts precisely and unambiguously.

## **ENTITIES AND RELATIONSHIPS**

The most important constructs within a logical data model are entities and relationships. An entity is a person, place, thing, or concept about which facts are to be recorded. Examples in Exhibit 1 are RENTER, RENTAL-AGREEMENT, PROPERTY, BEACH-PROPERTY, and MOUNTAIN-PROPERTY. A relationship is an association between two entities. Examples are RENTER rents RENTAL-AGREEMENT, PROPERTY is rented to RENTAL-AGREEMENT, and PROPERTY is of type BEACH-PROPERTY or MOUNTAIN-PROPERTY. In a logical data model diagram, entities are represented by boxes and relationships by arrows.

Entities have a number of properties. For instance, each entity has a name (such as RENTER) and a description (such as "person who obtains the privilege of residing on a specific property according to the terms of a rental agreement"). Entity sets (such as all renters) can be distinguished from entity occurrences (such as renter Harry Smith and renter Josephine Morgan). All occurrences within an entity set have the same attributes or detailed information items (such as NAME, ADDRESS, PHONE-NUMBER, and MAX-MONTHLY-RENT-AMT - all of which are attributes of RENTER).

## **ATTRIBUTES**

More formally, an attribute is a fact or nondecomposable (atomic) piece of information describing an entity. Nondecomposable means that an attribute represents the smallest unit of information that will be referenced at one time. For example, notice that a renter address is represented as one attribute (ADDRESS) but a property address is represented as two attributes (STREET-ADDRESS and TOWN-STATE-ADDRESS). This design enables the user to easily list all properties in a particular town and state;

listing only those renters in a particular town and state may be more difficult.

A particular entity occurrence can be identified by the values of its attributes. For example, renter Harry Smith can be distinguished from renter Josephine Morgan based on values of their attributes as illustrated in the sample value listing (sample entity occurrences with their respective attribute values) in Exhibit 2.

## EXHIBIT 2 RENTER SAMPLE VALUE LISTING

### RENTER

NAME (UIX)	ADDRESS	PHONE-NUMBER	MAX-MONTHLY -RENT-AMT
Harry Smith	12 Oak Lane, Springs	011-984-3158	400
Josephine Morgan	5 Central Ave, Cape Town	012-232-7990	650

Typically, all of the attribute values are not needed to identify a particular entity occurrence. For example, a particular RENTER can be identified by NAME (e.g., Harry Smith or Josephine Morgan). This identifying attribute or set of attributes (NAME in the case of the entity RENTER) is known as a primary key. The primary key of PROPERTY is the combination of two attributes, STREET-ADDRESS and TOWN-STATE-ADDRESS. Both attributes are needed to identify a particular property, because multiple properties may have the same value for STREET-ADDRESS (in different street addresses). RENTAL-AGREEMENT has a primary key consisting of three attributes: PROPERTY\*STREET-ADDRESS and PROPERTY\*TOWN-STATE-ADDRESS (identifying the property). Primary key attributes are written above a horizontal line in the logical data model



diagram (as in Exhibit 1) or on the far left above a (PK) designation in a sample value listing (as in Exhibit 2).

Some assumptions about definitions of the attributes are made when choosing the primary keys. These definitions should be explicitly understood and record within the design documentation or data dictionary. A data dictionary is a manual or automated repository of information about applications, data bases, logical data models, users, and access authorizations. Assumptions about the model are based on naming conventions used in the logical data model diagram without referring to the data dictionary. All entity, relationship, and attribute names are composed of English words, are frequently abbreviated, and are usually connected by hyphens. Abbreviations are used consistently. Moreover, each attribute name includes one word (the class word) that indicates the nature of the data represented by the attribute (e.g., NAME, ADDRESS, NUMBER, AMOUNT, DATE, FLAG).

Some of the attribute names include the name of another entity as a prefix, followed by an asterisk: for example, RENTER\*NAME, PROPERTY\*STREET-ADDRESS, PROPERTY\*TOWN-STATE-ADDRESS. These attributes are part of a foreign key. A foreign key is an attribute or set of attributes that completes a relationship by identifying the associated entity. The term foreign conveys the idea that the attribute belongs to or refers to another foreign entity. Thus, RENTER\*NAME in RENTAL-AGREEMENT identifies which RENTER is renting the agreement; and PROPERTY\*STREET-ADDRESS, PROPERTY\*TOWN-STATE-ADDRESS in RENTAL-AGREEMENT identifies which PROPERTY is being rented.

BEACH-PROPERTY and MOUNTAIN-PROPERTY share a special relationship with PROPERTY, designated by the squared-off arrow connecting the two to PROPERTY. BEACH-PROPERTY and MOUNTAIN-PROPERTY are each subtypes of PROPERTY, representing the same real-world object but having slightly different, more specific definitions and characteristics. They also contain the foreign key PROPERTY\*STREET-ADDRESS,

PROPERTY\*TOWN-STATE-ADDRESS, identifying with which PROPERTY each is associated.

Some of the arrows in the diagram are double-headed, and others are single-headed. The double-headed arrow represents a one-to-many relationship; for example, one RENTER may rent many RENTAL-AGREEMENTs, and one PROPERTY may be rented through many different RENTAL-AGREEMENTs (although not for the same week). A single-headed arrow represents a one-to-one relationship; for example, a PROPERTY is listed as a BEACH-PROPERTY or a MOUNTAIN-PROPERTY, but never as both.

A full logical data model consists of not only a diagram but specifications in a data dictionary. For example, although there may be multiple RENTAL-AGREEMENTs, for the same PROPERTY, they may not be for the same week; or, more generally, they may not be for overlapping time periods. As mentioned before, for two RENTAL-AGREEMENTs on the same PROPERTY (i.e., the same PROPERTY\*STREET-ADDRESS, PROPERTY\*TOWN-STATE-ADDRESS), the BEGIN-DATEs must be different (because PROPERTY\*STREET-ADDRESS, PROPERTY\*TOWN-STATE-ADDRESS and BEGIN-DATE constitute a primary key that uniquely identifies a RENTAL-AGREEMENT). However, a business rule involving END-DATE must also be specified. For example, when a new occurrence of RENTAL-AGREEMENT is inserted, the END-DATE must be later than the BEGIN-DATE. Moreover, the END-DATE must precede the BEGIN-DATEs of all existing RENTAL-AGREEMENTs having a BEGIN-DATE later than the new RENTAL-AGREEMENT's BEGIN-DATE, for RENTAL-AGREEMENTs for the same property. Otherwise, the rental agreement is inappropriate and does not make business sense. A data base implementation of the model should reject insertion of such an occurrence.

Such business rules are specifications that preserve the integrity of the logical data model by governing which values attributes may assume. True to the data-driven philosophy, business rules are identified without

consideration for exactly how the system will produce reports or enforce edit criteria. For now, the data and all of its relevant rules are analyzed independent of application requirements.

A logical data model, therefore, incorporates numerous rules about the integrity as well as the structure of information used within a business. Most of the rules conveyed by the logical data model diagram relate to structure. Other business rules specified within the data dictionary relate to integrity. For instance, the statement, "A RENTAL-AGREEMENT must be for a predefined PROPERTY but may be made by a RENTER who is not yet defined within the data base" is a key business rule (governing valid relationships between primary and foreign key attributes). "PHONE-NUMBER is a ten digit numeric attribute" is a domain business rule (governing types and ranges of values which attributes may assume). The example discussed involving END-DATE is actually a type of rule called a triggering operation (governing general effects of insert, update, or delete operations on other entities or attributes). Structure and integrity within the logical data model are equally important. It is useless to understand the structure of information within the business without understanding rules pertaining to that information's integrity, and vice versa.

## **2. THE ROLE OF A THREE-SCHEMA ARCHITECTURE**

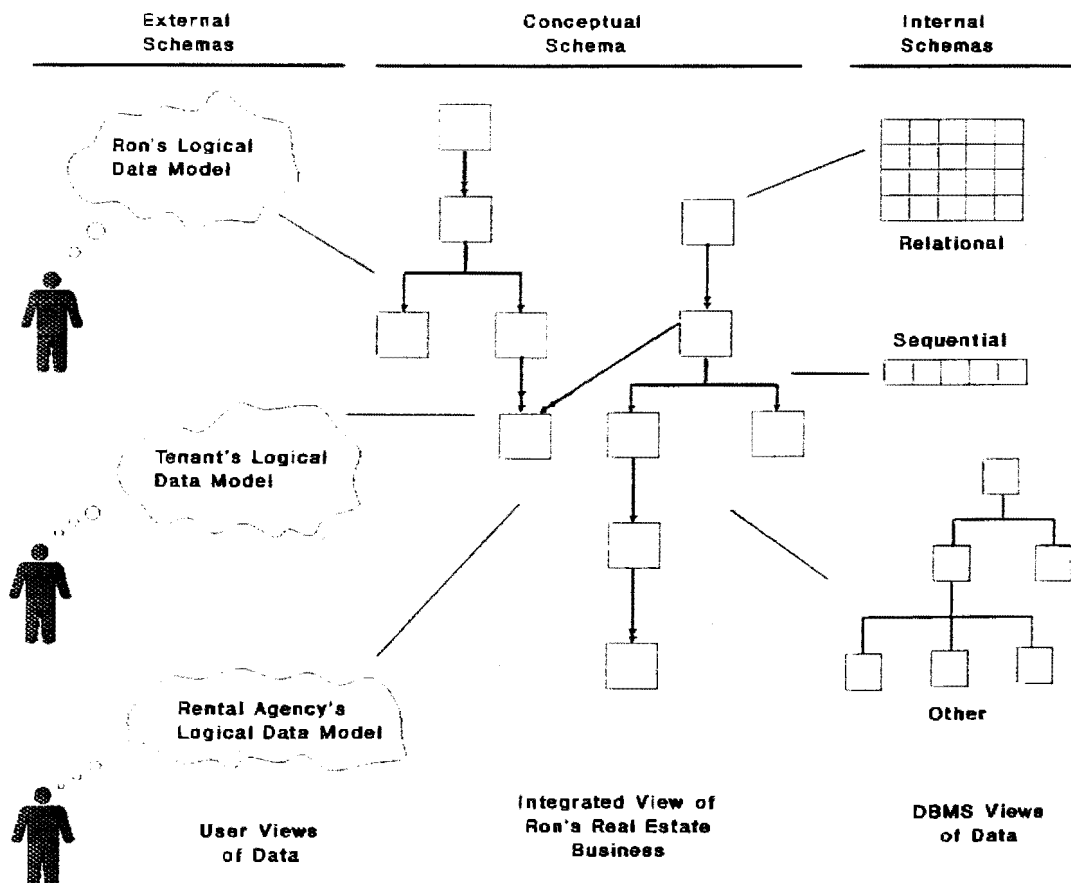
At this point, the reader may be questioning the relationship of a logical data model (supporting one area of the business or one set of business functions) to an integrated model of information used throughout the business. Such an integrated model is frequently called a conceptual model or conceptual schema and is discussed in the following paragraph.

In 1977, the American National Standards Institute (ANSI/X3/SPARC Committee) developed a set of requirements for effective data base management systems (DBMSs). These requirements were specified in terms of a three-part framework for organizing data managed by a DBMS:

- The external schema - Organisation of data as viewed by the user or programmer
- The internal schema - Data organisation as viewed by the DBMS's internal access logic
- The conceptual schema - An integrated view of all data used within the business

Exhibit 3 illustrates the mapping among external, internal, and conceptual schemas.

### EXHIBIT 3 ANSI/X3/SPARC THREE-SCHEMA ARCHITECTURE



The logical data modeling methodology described in this article is consistent with - and builds on - the ideas in the ANSI/X3/SPARC three-schema architecture. For example, the external schema for Ron's Real Estate Business is an organisation of data requirements from the perspective of Ron, the owner. Yet another external schema could be described from the perspective of a property tenant. That external schema might include information about only one rental property because the tenant cares about only the one property he is renting. It probably would include more details about that one property, such as days for garbage pickup and location of the nearest liquor store.

As the integrated view of all the business's data, the conceptual schema is a consolidation of all relevant logical data models. A conceptual schema can be developed as a prerequisite deliverable before implementing any data bases. Or, more practically, a conceptual schema can evolve by consolidating logical data models as they are built. The conceptual schema will help ensure consistency of multiple logical data models, defined to support different user groups or different areas of the business. Existence of one (even evolving) conceptual schema will also aid in designing shared data bases or internal schemas that support multiple external schemas. Finally, the conceptual schema will help ensure consistency across multiple data base implementations or internal schemas (e.g., making use of different technological strengths). Thus, a user can maintain one view (one external schema) of the data regardless of underlying implementations (internal schemas).

This article has generalized the ANSI/X3/SPARC definition of a three-schema architecture. ANSI/X3/SPARC defined the three types of schemas in terms of how they should be supported within a given DBMS (e.g., a relational data base product). The design principles of the three-schema framework are incorporated to produce logical data models and an integrated business conceptual schema that can be implemented through any (or, indeed, multiple) data base technologies. Thus, an architecture for DBMSs and also a specification for effective data base design practices are borrowed from

ANSI/X3/SPARC.

### **3. STEPS WITHIN LOGICAL DATA MODELING**

The following list is a set of criteria for an optimal logical data model:

- Structural validity - Consistency with the way the business defines and organises information
- Simplicity - Ease of understanding even by users of by nontechnical professionals
- Non-redundancy - Inclusion of no extraneous information; in particular, representation of any one piece of information exactly once (this may be a sub-criterion of simplicity)
- Shareability - Not specific to any particular application or technology; thereby usable by many
- Extensibility - Ability to evolve to support new requirements with minimal impact on existing base
- Integrity - Consistency with the way the business uses and manages information values

Satisfying these criteria requires more than assembling just any combination of boxes, arrows, and labels. Specifically, a set of rules and steps for applying those rules should be followed to ensure an optimal logical data model. Up to this point, this article has highlighted some of the rules, such as the rule that every entity should have a primary key and the rule that primary keys are unique. A sequence of steps for applying these rules is summarised as follows.

**Building Skeletal User Views.** Building a logical data model involves

examining one activity or business function at a time. The information required by each such function is individually modeled and the individual models are then integrated into one composite model. The model or representation of information required by one business function is known as a user view. Building a user view begins with the following steps:

- Identifying major entities (significant objects of interest) – Step LDM1
- Determining relationships between entities – Step LDM2

Adding Keys to User Views. The process continues with addition of key detail information items and the first, most important business rules. Specifically:

- Determining primary and alternate keys (identifying properties of entities) – Step LDM3
- Determining foreign keys (identifying properties of relationships) – Step LDM4
- Determining key business rules (rules that govern the effects of insert, delete, and update operations on relationships) – Step LDM5

Alternate keys are alternative choices of identifying attributes that (perhaps arbitrarily) are not chosen to be the primary key. For instance, if NAME is chosen as primary key for the RENTER entity in Ron's Real Estate Business, SOC-SEC-NMBR (Social Security number) might be an alternate key.

Key business rules define conditions under which primary and foreign keys may be inserted, deleted, or updated. One insert rule and one delete rule are established for each relationship. The insert rule determines valid conditions under which the foreign key may be inserted or updated in an entity occurrence. The delete rule determines valid conditions under which the primary key referenced by a foreign key may be deleted or updated.

These are the most frequently encountered and typically the most important constraints on insert, delete, and update operations. Because they define rules governing valid existence of particular entity occurrence, they are also called existence constraints.

**Adding Detail to User Views.** After identifying all key attributes (primary, alternate, and foreign keys), remaining non-key attributes must be added – step LDM6. Non-key attributes are the descriptive detail that users naturally associate with the entities. For example, ADDRESS, PHONE-NUMBER, and MAX-MONTHLY-RENT-AMT are non-key attributes in the RENTER entity. Each attribute is associated with the entity whose entire primary key is required to identify it. (The primary key may be a set of attributes).

**Validating User Views Through Normalisation.** Step LDM6 relies primarily on an intuitive process of associating attributes with seemingly the proper entities. In step LDM7, that process is checked through a more structured, formal technique: validating normalisation rules – step LDM7.

Normalisation is a body of theory addressing analysis and decomposition of data structures into a new set of data structures exhibiting more desirable properties. Specifically, normalisation increases the certainty of achieving an optimal logical data model. Each entity and its associated attributes are successively examined for structural redundancies or inconsistencies. Such redundancies or inconsistencies are due to assignment of an attribute to the wrong entity (e.g., association of customer credit rating with the ORDER rather than with the CUSTOMER entity). Such problems are eliminated by reassigning attributes to more appropriate entities or, in some cases, by decomposing entities into smaller, simpler entities. The result is a model that is at least structurally consistent.

Normalisation does not ensure that the model correctly reflects the business meaning of data requirements. Thus, normalisation should be employed as a refinement technique (step LDM7) only after completing a thorough



business analysis using the techniques in step LDM1 through LDM6.

**Determining Additional Attribute Business Rules.** Normalisation also does not fully address business rules. These should be part of a logical data model to ensure that not only the structure but the values of the data correctly reflect business operations. Steps LDM3 through LDM5 already have uncovered the most significant business rules – those governing integrity of primary and foreign key attributes. At this point, two additional types of attribute business rules can be identified by:

- Determining domains (constraints on valid values that attributes may assume) – Step LDM8
- Determining triggering operations (rules that govern the effects of insert, delete, and update operations on other entities or other attributes within the same entity) – Step LDM9

The term domain includes data type, length, format or mask, uniqueness of values, null support (whether a value must be present), allowable values, default value if applicable, and business meaning. Domains verify whether values assigned to an attribute make business sense. They also help in determining when it makes sense to compare or manipulate values of two different attributes.

The term triggering operation refers to the most generalised form of business rule, encompassing domains and key business rules as well as other attribute business rules. They reflect the user's understanding of all rules that make some sets of data values correct and others incorrect in the business world (e.g., for a given RENTAL-AGREEMENT, END-DATE must be later than BEGIN-DATE).

The intent in defining business rules is to clarify all data-driven constraints on the data values. In other words, those rules that always hold true are defined, regardless of any particular processing requirements. Because

these rules are application-independent, they are defined as part of the data design rather than as part of the application design. If instead they are treated as part of the application, they would have to be completely, consistently, and redundantly specified as part of every application accessing the data.

**Integrating User Views.** The final logical design steps combine user views into one consolidated logical data model:

- Combining user views into one model - Step LDM10
- Integrating with existing data models - Step LDM11
- Analysing for stability and growth - Step LDM12

In Step LDM10, user views defined for different business functions are combined; or perhaps different user groups are combined into one model. For example, one user view for Ron's Real Estate Business may be defined from the perspective of Ron, the owner, as illustrated in Exhibit 1. If Ron also uses rental agencies to assist him in locating potential renters, a different user view can be developed from the perspective of Clancey's Rental Agency, dealing with multiple property owners and additional rental properties. Thus, Clancey's user view includes additional entities, subtypes, relationships, attributes, and business rules. Presumably, in some areas, Clancey's and Ron's user views overlap; in some areas, they may even conflict. In step LDM10, the overlaps are consolidated, any inconsistencies resolved, and new, inter-view relationships and business rules added to form one composite logical data model.

In step LDM11, this consolidated logical data model is examined in light of models developed for other purposes. Again, overlaps and some inconsistencies will be discovered. Most of these other models may already have been implemented as data bases; thus, the models may not be able to be changed if errors or omissions are discovered. The objective here is to understand and document relationships (including inconsistencies) among the designs.

This can be accomplished by comparing and defining mapping among the logical data models, specifically through mapping to the business conceptual schema. The conceptual schema evolves by combining the logical data models, merging them two at a time, similar to merging user views. Then mapping are identified between each logical data model and the business conceptual schema, including:

- Differences in names
- Operations performed on the conceptual schema to obtain constructs within a particular logical data model
- Interrelation of business rules

A business conceptual schema allows multiple logical data models and multiple data base implementations that are consistent with one another and with the business's overall operations as well as being representative of individual user perspectives.

Finally, in step LDM12, consideration should be given to future business changes that may affect the current logical data model. Those that are significant, imminent, or probable are incorporated into, or at least documented with, the logical data model. The goal is to maximize stability of the logical data model – to ensure that correctness and usefulness will necessitate few changes over a reasonable period. It is most important to incorporate any changes that affect the business conceptual schema, because these are likely to have major influence on one or more individual logical data models.

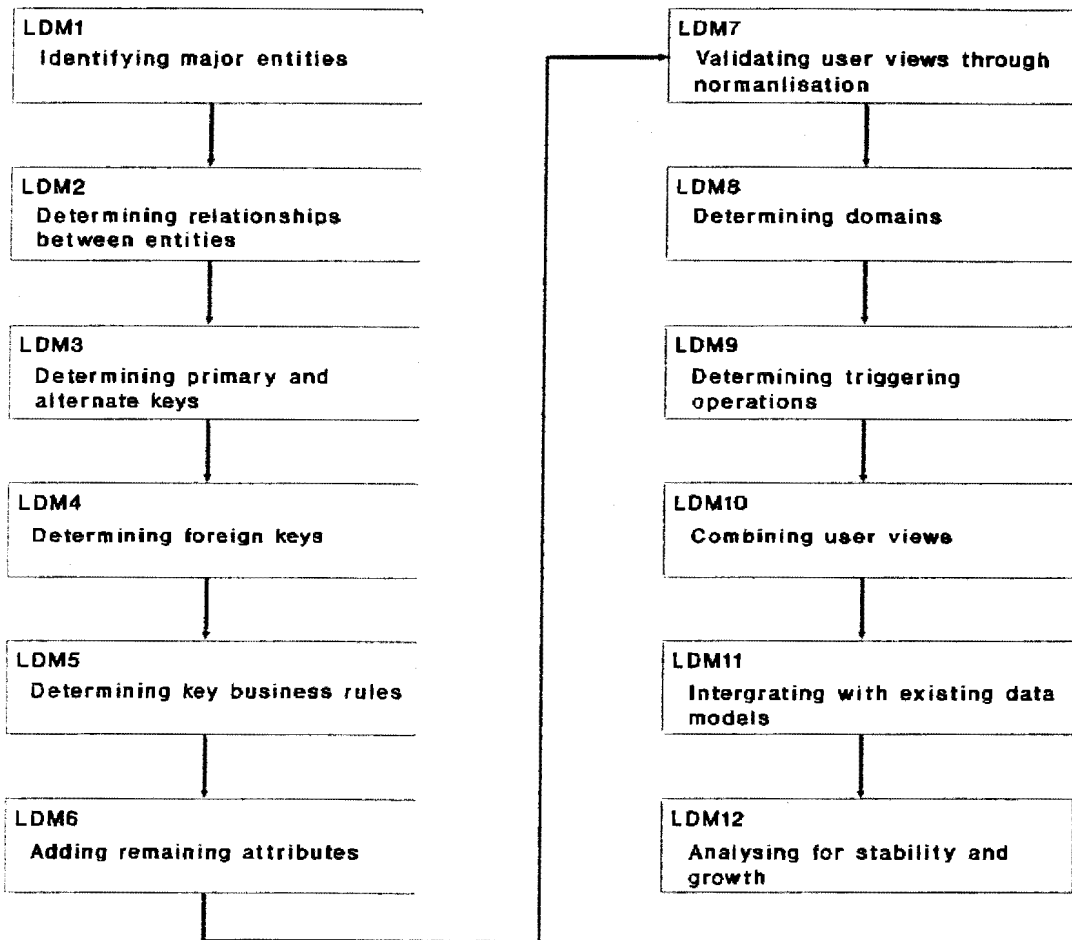
Exhibit 4 illustrates all the steps in the logical data modeling process. The result is a simple, clear, shareable, stable model exhibiting minimal redundancy and accurately reflecting both the structure and the integrity of information used within the business.

#### **4. CRITICAL SUCCESS FACTORS IN LOGICAL DATA MODELING**

For successful application of a logical modeling methodology the manager should:

- Work interactively with the users as much as possible – Building confidence that the manager and the users share an understanding of the information requirements
- Follow a structured methodology throughout the logical data modeling process – Choosing a methodology before beginning and stay with it
- Employ a data-driven approach – Building a logical data model that represents how the business uses and manages information, unbiased by any particular processing requirements or technological considerations.
- Incorporate both structural and integrity considerations in the logical data model – Addressing structure alone might reveal how to organise records within a data base but will not ensure that the data values accurately model business operations
- Combine both conceptualisation and normalisation techniques into the logical data modeling methodology – Conceptualisation identifies entities, relationships, and keys (steps LDM1 through LDM6) and ensures that the model on paper truly represents the use of information within the business. Normalisation (step LDM7) ensures that the model is structurally consistent and logical, and has minimal syntactic redundancy
- Use diagrams to represent as much of the logical data model as possible – Diagrams are clearer, simpler, and more concise than text

## EXHIBIT 4 THE LOGICAL DATA MODELING PROCESS



- Build a data dictionary to supplement the logical data model diagrams – Diagrams cannot convey everything. Textual specifications of entity/relationship/attribute definitions and business rules are also needed

### 5. BENEFITS AND APPLICABILITY OF LOGICAL DATA MODELS

Logical data modeling is a critical prerequisite to effective data base design. Logical data modeling can also contribute to the success of other less obvious endeavour, including:

- Assessing technological feasibility – A data-driven logical data model

represents data structures, relationships, and rules without any compromises to accommodate technological limitations. The degree to which a particular technological implementation can materialise all components of the logical data model provides an effective measure of the technology's optimality for related applications

- Assessing software packages – In particular, the degree to which the data base design of an application software package approximates the logical data model serves as one gauge of the package's ability to meet business requirements
- Assessing the impact of future changes – A logical data model provides a clear picture of the underlying data base design, unobliterated by technological detail. It thus aids significantly in evaluating the effects of changes in business requirements or technology
- Facilitating end-user access to data – A logical data model is a business representation of information structures and rules, and can be understood by users. It can provide these users with an intelligible map of their data within an end-user computing environment
- Facilitating a strategic overview of the business's information needs – The logical data modeling process, minus some of its more detailed steps, can be applied to build a strategic data model or high-level representation of all the business's information requirements. A strategic data model can be thought of as a summarised version of the business conceptual schema. The strategic data model can assist in analysing information interrelationships throughout the business. It therefore can contribute to more effective planning of data base and application implementations. For example, it can aid in planning and relating detailed operational systems (those required to run the daily business) and summarised, executive decision support systems
- Migrating data from one technology to another – Data bases can more

easily migrate from one technology to a newer or more appropriate technology when a current logical data model exists. The model can be translated into a new data base implementation. The logical data model can also be used to design extract procedures. If a logical data model does not exist, one should be built by treating the existing data base as a user view and by working with users to discard unnecessary information requirements as well as add new ones

## 6. SUMMARY

Logical data modeling is the first phase of an effective data base design process. It is a technique for understanding and capturing business information requirements that must be met through a data base implementation.

The logical data modeling process begins with the definition of a user view or representation of information requirements for one business function, user, or user group. It culminates in the integration of user views into one composite logical data model. Multiple logical data models are related through their mapping to a business conceptual schema, an integrated logical data model representing all the business's data requirements at their most detailed level.

Development of logical data models is a critical component of a data-driven design philosophy. Success stories of data base designs that were founded on logical data models and business conceptual schemas are increasing. Recently, emergence of computer-aided software engineering (CASE) tools has made possible automation of portions of the logical data modeling process. In addition, the dramatic surge and acceptance of commercial relational DBMS products have provided opportunities for designing data bases that closely resemble logical data models. For all of these reasons, the benefits of a stable logical data model are perhaps more widely understood and more easily accomplished than ever before.

Source: Fleming, C.C. & Von Halle, B. 1990. Logical data modeling. In Auerbach data base management. 22-30-21. New York:Auerbach: 1-14. (Acknowledgement : Handbook of relational data base design. 1989. Addison:Wesley Publishing Company Incorporated.)



## GLOSSARY OF DATA TERMINOLOGY

This glossary describes as simply as possible some of the basic concepts of data management.

### **Attributes (data elements) and attribute types**

An attribute is described in the Concise Oxford Dictionary as "a quality ascribed to anything". In data modelling, an attribute is used to describe the characteristics of the entities. An attribute is a descriptive value or property associated with an individual entity. For a building, an attribute may be an address – 12 Church Street, for example. An attribute type is the set of all attributes of a particular type. In the above example, the attribute type would be street address. An attribute appears in a data base as a field or data item, and the attribute type appear as the field name or data-item name.

See also **data element** and **data item**.

### **Batch environment**

A sequentially dominated mode of processing; in a batch, input is collected and batched for future processing throughout the day. Once collected, the batch input is transacted sequentially against one or more data bases.

### **Canonical model**

A data model that represents the inherent structure of data without regard to either individual use or hardware or software implementation.

### **Conceptual schema**

A consistent collection of data structures expressing the data needs of the organisation. This schema is a comprehensive, base-level and logical description of the environment in which an organisation exists, free of physical structure and application system considerations.

**Data**

Data is a description of something concrete, a situation or a concept in a speakable or recordable medium. (For example the alphabet, use of language or recording.) Data represents raw facts and figures that are normally meaningless by themselves. Data can also be described as a recording of facts, concepts, or instructions on a storage medium for communication, retrieval and processing by automatic means and presentation as information that is understandable to human beings.

**Data or entity model**

An entity model (entity-relationship) or data model, is a representation of the organisation's business data in terms of its entity groups, entities data groups and attributes and their relationships.

An entity model (also known as a conceptual data architecture) represents the data foundation of the organisation's business. It is important to note that data modelling is concerned with building an accurate picture of the real world that can be used to develop information systems that track the status of real world events and answer a range of questions about them. A data model is also known as a Bachman diagram.

See also **entity-relationship model**. See also **corporate entity model**.

**Data analysis**

The activity of identifying the entities that are of interest to an organisation and describing them, their attributes, and their relationships in a data or entity model.

**Data architecture**

A data architecture comprises the high-level corporate entity model, (enterprise model) which is subdivided into business entity models corresponding to important business activities. The business entity models are normally further refined to produce the overall logical data models.

**Data base**

A collection of interrelated data stored (with controlled redundancy) according to a schema. One data base can serve one or more applications.

\* **Data base management system (DBMS)**

A set of software facilities to create, access and control a database.

**Data base technology**

The technology that supports the data base which includes the Data Base Management System (DBMS). Corporate data base technology refers to the data base technology that supports the corporation or organisation as a whole.

**Data definition**

The specification of the data entities, their attributes, and their relationships in a coherent data base structure to create a schema.

**Data definition language (DDL) (also called a data description)**

The language used to define the data base schema.

**Data description language (DDL) (also called a data definition language)**

A language for describing data.

**Data dictionary system**

A software tool that allows the recording, storing, and processing of such meta data as data definitions, descriptions, and relationships between programs, data, and users.

See also **information resource dictionary system** and **data encyclopedia (data dictionary)**.

**Data element**

(1) An attribute of an entity.

(2) A uniquely named and well-defined category of data that consists of data

items and that is included in a record of an activity.

See also **data item**. See also **attribute**.

### **Data independence**

The property of being able to modify the overall logical or physical structure of the data without changing the application program's view of the data.

### **Data encyclopedia**

A data encyclopedia holds information about entities (and their attributes) and the processes that use them. At the programming level, an encyclopedia stores the names of data items, their formats, and the rules governing their content. At the database level, it stores the descriptions of the logical and physical database structure. The information held in a data encyclopedia provides support for all stages of the system development life cycle, from the initial analysis to applications testing and maintenance.

#### **\* Data directory**

The most basic form of a data encyclopedia is a data directory. A directory holds descriptions of the data used by the database management system, and it is used solely by the database administrator to manage the database.

#### **\* Data dictionary**

It is a lower level encyclopedia system that contains information about the applications that use the data (programs, screen layouts, report layouts, and so on). Recording this information in a data dictionary allows the effects of changes to applications or data to be assessed easily.

#### **\* Data repository**

The encyclopedia system in its broadest sense is known as a repository. It contains information about data and applications and supports the whole of the development life cycle, together with information about some

operations activities such as change management, operating-systems management and network management. A data repository is thus a large and complex database that collects data from several data-management systems and from other software systems.

**Data group**

A data group is a collection of related data attributes (elements) and describes something concrete, conceptual or a situation. It is a logic grouping of data.

**Data item**

A discrete representation having the properties that define the data element to which it belongs.

See also **data element**.

**Data-management systems**

Software facilities for the management of data. These include data dictionaries (encyclopedias), database management systems, and distributed database management systems.

**Data base management system (DBMS)**

See also **data management systems**.

**Data dictionaries**

See also **data encyclopedia**.

**Data manipulation language (DML)**

- (1) A programming language that is supported by a DBMS and used to access a data base.
- (2) Language constructs added to a higher-order language (for example COBOL) for the purpose of data base manipulation.

**Data modelling**

The process of creating logical representations of data. The logical data model reflects the fundamental data needs of the organisation, not the current organisation of that data. The data model is independent of the processing system chosen.

**Data record**

An identifiable set of data values treated as a unit, an occurrence of a schema record in a data base, or a collection of atomic data items describing a specific object, event, or tuple (for example, row of a table).

**Data security**

The protection of the data in the data base against unauthorised disclosure, alteration, or destruction.

**Data structure**

A logical relationship among data elements that is designed to support specific data manipulation functions (for example, trees, lists, and tables).

**Data type**

The definition of a set of representable values that is primitive and without logical subdivision.

**Directory**

A table giving the relationships between items of data. Sometimes a table (index) giving the addresses of data.

See also **data encyclopedia (data directory)**.

**Domain constraints**

The constraints that define the type, length, format, and allowable values for individual data items (attributes).

### **Entities and entity types**

An entity is described in the Concise Oxford Dictionary as "a thing's existence as opposed to its qualities or relations; a thing that has real existence." An entity is represented by data. An entity is a person, place, thing, or event, represented by a noun, about which data relevant to the organisation's business is to be maintained. Sets of classes of entities are called entity types. An entity type covers all the entities relevant to the enterprise that fit into a given definition. Thus, the entity "Barlows" is an instance of the entity type "company". In a data base, entities usually appear as individual records or relations, and types appear as the record name.

An entity group is a logical group of entities, also represented by a noun, that related to a major functional area of interest within the organisation's business. It is synonymous with the term subject data base.

### **Corporate entity model**

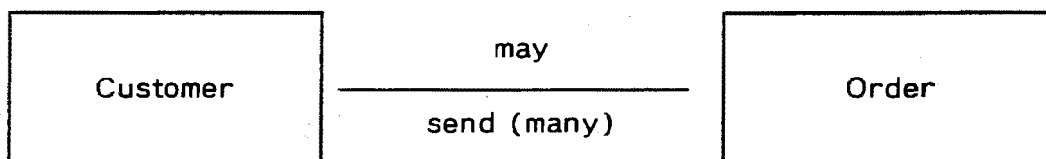
A high-level entity (data) model of the major business entities in an organisation.

### **Enterprise**

Relating to the whole organisation or corporation.

### **Entity relationship model**

A chart showing the relationship between entities and in particular whether the relationship is one to one, one to many or many to many. For example, a one to many relationship would be shown as:



The process is used to help define data relationships as part of the data base design. It can also be viewed as the logical data structures, including operations and constraints, provided by a

DBMS for effective data base processing.

**External schema**

A logical description of a user's method of organising data. Some attributes or relationships can be omitted from the corresponding conceptual schema or can be renamed or otherwise transformed.

See also **view and subschema**.

**Field**

See also **data item**.

**Flat file**

A collection of records containing no data aggregates, nested repeating data items, or groups of data items.

**Hierarchical model**

A data model providing a tree structure for relating data elements. Each node of the tree corresponds to a group of data elements of a record type and can have only one superior node or parent.

**Information engineering**

Information engineering involves the application of structured techniques not just to one system, but to the whole enterprise. The emphasis is on the information needs to be stored and maintained. This leads to a new front-end phase of the development cycle: "Strategic Planning" that focuses on corporate goals and business needs.

**Information resource dictionary system**

An information system that contains data about another (for example, target) information system, its environment, and its implementation – possibly on a DBMS.

See also **data dictionary system**.



**Integrity**

The term integrity is used to describe the measures taken to ensure the correctness of a database. These measures ensure that the database structure always reflects the structure of the data model used to design it, and that the data it contains conforms to the descriptions in the data dictionary. It can be defined as the completeness, approval, accuracy and consistency of data.

**Internal schema**

The schema that describes logical structures of the data and the physical media to define the physical storage.

**Item**

See also **data item**.

**Key**

A data item or combination of data items used to identify or locate a record instance (or other data grouping).

**Meta-data**

Data describing data – as does a data dictionary.

**Network model**

A data model that provides data relationships on the basis of records and groups it records (for example, sets) in which one record is designated as the set owner, and a single member record can belong to one or more sets.

**Normalisation**

A data model and a relational database should contain data that has been "normalised". The process of normalisation occurs during data analysis and is used to rationalise data into its simplest form. In this form, each data item contains only one fact, and each data item is directly related to the data group (record) it is in. Normalised data can be handled in a much more flexible way than unnormalised data.

### **Natural forms**

- \* **First normal form** Data that has been organised into two-dimensional flat files without repeating groups.
- \* **Second normal form** Data that functionally depends on the entire candidate key.
- \* **Third normal form** Data that has had all transitive dependencies on other data items within the record removed, except for the candidate key.
- \* **Fourth normal form** Data whose candidate key is related to all data items in the record and that contains no more than one non-trivial multi valued dependency on the candidate key.

### **Operating system**

Software that enables a computer to supervise its own operations and automatically call in programs, routines, language, and data as needed for continuous throughput of different types of jobs.

### **Physical representation**

- (1) The representation and storage of data on such media as magnetic disks.
- (2) The description of data that depends on such physical factors as length of elements, records, and pointers.

### **Record**

An aggregation of values of data items or elements.

### **Referential integrity constraints/business constraints**

#### \* **Referential constraints**

The constraints that ensure the integrity of references between the rows of one table and the rows of another table (or several tables).

#### \* **Business constraints**

Constraints that ensure the integrity of a data item value in one table, given one or more data item values in the same table or in other tables.

**Relational model**

A data model allowing the expression of relationships among data elements as mathematical relations. The relation is a table of data representing occurrences of the relationship as the rows (for example, tuples).

**Relationship**

A relationship indicates the dependency in terms of right of existence between entities, entity-groups and data groups in a specific system.

**Repository**

The database that contains the data used in the development and maintenance cycle.

See also **data repository**.

**Schema**

A complete description of the data base in terms of the data characteristics and the implicit and explicit relationships between data types.

**Security**

The protection provided to prevent unauthorised or accidental access, updating, copying, removal, or destruction of the data base, or its elements.

**Subschema (external schema or user schema)**

A representation of a part of the data base as used in a specific application or by a particular user group.

**Transaction**

A command, message, or input record that explicitly or implicitly calls for a processing action (for example, updating a file). A transaction is atomic with respect to recovery and concurrency.

**User**

A person or process issuing commands and messages to the information system.

**View**

An external relation that consists of attributes retrieved or derived from one or more base relations joined and projected as given in the view definition.

See also **external schema**.